Global Journal of Engineering Innovations & Interdisciplinary Research

Original Article

Human Computer Interaction Through Eye Ball

S Ramana Reddy [1], M Prashanth[2], G RakeshYadav[2], M Jayanth Reddy[2]

[1]Assistant Professor, Department of Artificial Intelligence and Data Science, Vignan Institute of Technology and Science, Hyderabad, India

[2]UG Student, Department of AI&DS, Vignan Institute of Technology and Science, Hyderabad, India

Correspondence

T. Sai Lalith Prasad

Assistant Professor , Department of Artificial Intelligence and Data Science, Vignan Institute of Technology and Science, Hyderabad, India

- Received Date: 25 May 2025
- Accepted Date: 15 June 2025
- Publication Date: 27 June 2025

Keywords

OpenCV; MediaPipe; PyAutoGUI; Face Mesh; Speech To Text ; CMU Sphinx

Abstract

Somepeoplehavehealth-relatedconditionsthatimpedetheir utilization of computers. In that regard, it is quite necessary to come up with redesigns for effective use of computers that meet the requirement of people living with disability. The human eye can be considered as a promising replacement of standard input devices. This paper presents a web camera as an option with which a person would control the cursor by moving the eye. For the pupil detection, we use OpenCV, so the eye frames obtained would be sequenced for the purpose of controlling the movement ofthemouseonthescreen. Also, Eye Aspect Ratio (EAR) is computed for determining the gaze movement, i.e., lateral movement of the left or right eye using the OpenCV library under Python programming. This particular method aims to optimize the computer environment in the very focus of those with physical challenges by providing them with a way to control the computer without using for instance a mouse.

Introduction

In the last couple of years, I have witnessed the revolutionary growth of computer assistive technologies aimed at facilitating the use of computers among the physically challenged. Conventional devices such as keyboard and mouse may pose problems or are out of reach for many physically challenged people, and this limits their capability to interact with the digital world. But the emergence of technologies like eye-controlled computing can help to change this scenario.

With the development of such technologies, it is now possible for people to operate a computer just by looking at the screen. Even the basic functions such as cursor movement, click, and scroll are performed by a glance, allowing people with physical impairments to use computers more easily. This technology has the potential to greatly enhance the quality of life for many disabled people [1].

The most fascinating applications of eye-controlled computing are those that bring intuitive and natural human-computer interaction with technology into the realm of those who have any physical impairment. Surf the net, open Word documents, or even play games with your eyes! An ideal system similarly targets all sorts of disabled individuals as well as those who want to surf the Internet or such other activity but find it difficult to go through a mouse or keyboard because of a condition or at any time.

Here, we describe a mouse pointer control program for gaze-induced activities using off-the-shelf web cameras along with some computer vision tricks. The system uses OpenCV and MediaPipe for tracking eye landmarks; this allows the user to operate the computer through slight movements of the eyeball or even blinking [2].

The concept driving the development of this technology is called Eye Aspect Ratios or EAR, which is a wonderful measure that determines whether the eyes are open or closed. With this, thesystemcan detect andinterpret even thesmallest movements visible on the eyes: for instance, blinking can actasaclick,orgazingin aspecified directioncan shift thecursor across a screen. The system would understand such minute changes and, as magic, translate them into actual activities on thescreens. Result?Thisensures smooth, intuitive experiences in which the computer seems to be responding directlyto your intentions, and even the more complex tasks are easy to control. [3]

As already said, the most paramount issue regarding eye- tracking technology is that it should be applicable in different lighting conditions and for various eye shapes. Eye-tracking cameras must be sensitive enough to detect eye movement in almost any environment, which is not easy; one solution offered by the system is intelligent image processingtechniques to enhance the contrast of the eye region tofacilitate detecting eye movements under tricky lighting conditions [4].

Hence, while the camera continuously delivers real-time video feed, it analyzes eye landmarks through time to trigger movements

Citation: Reddy SR, Prashanth M, Yadav GR, Reddy MJ. Human Computer Interaction Through Eye Ball. GJEIIR. 2025;5(4):073.

such as clicks left and right or scrolling. It also becomes very interactive since it becomes quite like using a mouse with an eye [5]. It is now integrated with vocal recognition to even make it easier. For instance, opening an app,dictating something, or controllingvolumecould be made possible by voice. And combined with eye-gesture commands, the system becomes more powerful because it can give such freedom to users in totally not touching the computer. It is a total breakthrough for the mobilities impaired because itcomes in that genuinely hands-free interface with technology [6]. It is a highly adaptable system which learns and modifies itself according to the unique eye movement pattern of the individual user and becomes more precise and personalized over time in usage. It improves the more it is used, so that every person is getting the most seamless experience possible. Furthermore, because it features real-time feedback, itbecomes more like an extension of the human body and feels immediate to each eye movement [7]. Be it a person sitting at home in a study or sitting in a clinic, the system understands the needs of a person and delivers a smooth experience. Dynamic filtering techniques are also at work to ensure precise tracking of eye movements even under adverse lighting conditions. So whether it's a dimly lit room or bright fluorescentlights, thesystem keeps working. Havinglarge-eye with small movements brings precise and various capabilities to the system—reading documents or browsing websites [8].

This system is so fast that when it comes to responsiveness, it's almost classless. Real-time processing creates no delay at all between eye movement and processes on the screen, givingmuch smoother performance with reduced frustrations. This is highly of use to users depending on the system to carry out their everyday activities. Now, they can do it much faster and at the same pace as using a mouse [9]. The machine learning algorithms will even learn very small eye movements, translating these into cursor movements on screen; thus, the user will have control over the cursor no matter how much effort is used.

Additionally, the complete system is customized. Users can set the sensitivity of the eye-tracking feature, whether they want the micromanaging fine-tuning of movements for delicate tasks or fast operation responding for gross movements. The flexibilityofthe system is such thatit can learnand then recognize favorite gestures like specific bangs or areas for invoking specificactions [10]. This way, the end-users will be really personalized, with experiences that are really tailored to meet their requirements.

The long-term aim of this technology is simply to make computers available for those who are physically disabled. Eye contact system is such that it ends up removing traditional input systems, allowing those who struggle with some sort of ability deprivation to use it for whatever it is that they may want to do-online possible working, learning, or playing. For everyone, this means having a much more extensive experience with computers for effective and independent operation.Thisis very much in the direction of using eye-tracking, voice recognition, and real-time processing all together to take steps toward creating a culture in which society has as much exposure to technology as possible.

Methodology

The system was with the mind as it was executing seamless processes which were hands-free and integrated with advanced eye-tracking and voice recognition technologies. It makes use of a standard webcam that captures the user's eye movements, turning them into the user's eye movement pattern — thus, it "x-rays" the state. It uses eye movements in the form of, e.g., cursor movement, clicks, and scrolling to make the process simple for the user to control a computer intuitively.

The system is run by strong libraries like OpenCV and MediaPipe that are programmed in such a way to identify the user by the eye landmarks to ensure that the control is reliable and spot-on. The use of voice recognition in the setup also elevates the system because it gives the users another possibility to do things such as opening applications, dictating text, or adjusting volume. Thus, the system becomes more accessible since voice recognition is additionally included.

The system is no longer just limited to hand gestures and voice commands but also integrates other functions as well. Through the system, the users' interactions result in its continuous learning and improvement due to its ability to be constantly in the learning process, becoming more proactive in responding correctly to the user's needs. This feature is the very one that secures unique user care, which in turn is the base that allows the system to become more and more efficient over time, thus making such an interaction tool very convenient for those with physical limitations to navigate the digital world safely, faster, and with a greater degree of independence.

On top of that, the compatibility of this system with different environmental setups increases its role in the facilitation of a real-life technological landscape where everyone is included.

Implementation

This entire procedure for using mouse movement by both pupil and mouth movement was achieved as follows:.

Face Detection

Face detection is done with the help of MediaPipe, Face Mesh Model, which serves to give real-time face landmark detection in a real-time basis. It captures live video frames through the webcam that are simply converted to RGB format for operating efficiency. Frames are processed to extract up to 468 significant facial coordinatescorresponding to facial features, such as the eyes, nose, mouth, and jawline. The location of the face in the image is known by detecting these facial points. Using landmarks, movements are detected in the face, including eye gesture detection and blink detection, to carry out such tasks, etc., cursor movement and click action. This method of face detection is the basis for a hands-free-computer control system that enables users to operate the machine througheye movements and facial movements. The use ofMediaPipe Face Mesh guarantees accurate real-timetracking of facial landmarks, essential in making the eye- controlled system function in different environments. The Face mesh visually summarized in Figure 1, highlighting each step and its contribution to the final output.

Eye Detection

The MediaPipe Face Mesh makes it possible to get features for the eye detection which are very accurate facial landmarks. These also have eye-specific points. The video frames arecapturedas live frames andthenprocessed bythe system to find out the key eye landmarks related to the eyelids and iris. Then the two eye landmarks' vertical and horizontal distances are used to calculate the Eye Aspect Ratio (EAR), which detects whether the eyes are opened or closed. This analysis is helpful for determining eye movement and blinking, which will then be translatedactions such as mouse clicks or cursor movement on the computer screen. Real-time tracking through eye landmark

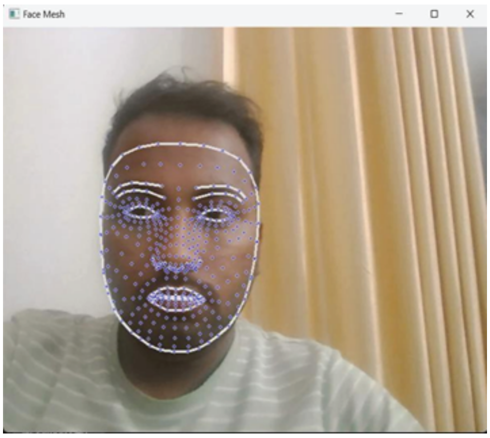

Figure 1: *Facemesh*

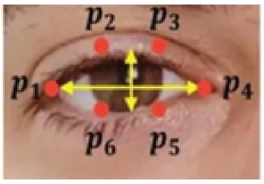

Figure 2: Openeye

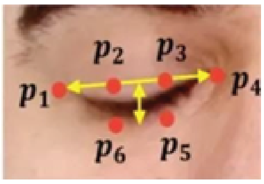

Figure 3: Closed eye

allows the system to make a computer hands-free for the individuals having physical disabilities.

Mouth Detection

Mouth detection is done using the specific landmarks around the mouth region captured by the MediaPipe Face Mesh. Once the video frames are captured and processed in this system, it will then be analyzing the points that correspond to the upper and lower lips and measuring the distance between them to tell whether the mouth is open or closed. As soon as the distance between the upper and lower lip touches a certain threshold, that means the mouth is opened. This feature can be used for triggering events such as voice recognition for hands-free commands or gesture-based activation of the system. The ability to detect the status of the mouth in real-time, enables smooth interactions, not to mention providing accessibility to users with physical impairments.

Voice to Text

The CMU Sphinx capability converts the oral word into peculiar written text and makes it possible for a person to operate the computer without resorting to other input devices. This means that it is possible to employthe system with the hands-free types of operations like opening applications,

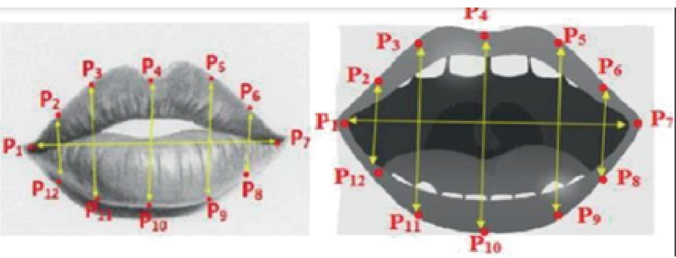

Figure 4: *Open and closed mouth atio*

entering text, oradjusting settings by real-time voice commands captured through a microphone. It learns to become more accurate as time goes and becomes more familiar with different speech patterns and environment conditions, along with noise filtering feature to minimize outside sounds. It goes further than typing commands in that it can also command volume switching or program initiation, all through voice commands. It ministers with the assistive technologies used, such as eye tracking, to provide a more comprehensive and complete solution to enable mobility- impaired people as greater access to their digital world independently.

The above systems apply CMU Sphinx. A spoken word is translated into written word through this system and enables users to command their computers without the conventional input devices. It also enables a mean of typing, opening applications, adjusting settings-much hands-free-by real-time voice commands. The common application of hands-free voice command control includes: typing; opening specific applications- installed in the computer-or performing certain actions such as adjusting settings. Over time b will learn to become even more accurate with speech patterns, locations, and noise filtering. The system itself goes beyond typing, and it can switch the volumeon or start a program by means of voice commands. Other assistive technologies like those used in eye tracking make this even more accessible since it gives a whole solution to persons with mobility impairments to interact with their digital worlds more.

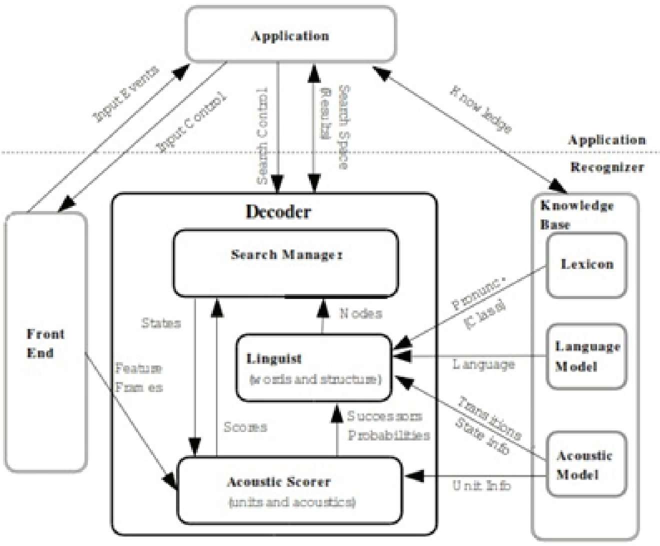

Figure 5: CMU Sphinx Architecture

Scrolling

Scrolling according to head movement is the part where the user is required to move their head to interact with the system in a quite intuitive and naturalistic manner. Thishead scrolling works by computer vision trackingalgorithms that pick specific face constraints, specificallythe tip of the nose. If the nose has moved up or down, the program understands that the person's head tilts in that respective direction, prompting scrolling action by up or down scrolling, respectively. And, depending upon the agitation the head has caused, the speed of scrolling will automatically change, hence giving the very natural feeling and more comfortable experience.

This is very convenient for hands free communication, especially for people who have little mobility left in them and can't handle typical input devices. The new freedom for the user features can access the digital touch, probably reading documents, browsing through images, or moving down to long web pages. This mode of interaction be avoided even without the use of physical touch, reduced strain, and ergonomic communication.

The currently described head-scrolling mechanism caneasily be embedded into an inclusive system, such as eye- tracking and voice recognition mechanisms, thus enabling enhanced overall aid accessibility. Then, with the assistance of such additional technologies, various actions could be performed, such as moving a mouse cursor, dictating text to a computer, and executing voice commands. This cross- integration createsaninclusiveuser experienceandprovides those with disabilities a better digital experience that supports their autonomy and efficiency in human-computer interaction.

Algorithm

- The system initializes the camera and libraries for facial landmark detection and voice recognition.
- It captures video frames and processes them for face detection using Mediapipe.
- Eye and mouth gestures are tracked to trigger scrolling, clicks, and voice-to-text activation.
- Scrolling is controlled by head movements, with vertical nose movement determining direction.
- Speech is captured when the mouth is open and converted to text using CMU Sphinx.
- The mouse cursor is controlled by eye landmarks, with horizontal and vertical movements.
- The system continuously processes video frames, checking for gestures and voice commands.
- The process terminates when the 'q' key is pressed, releasing camera and resources.

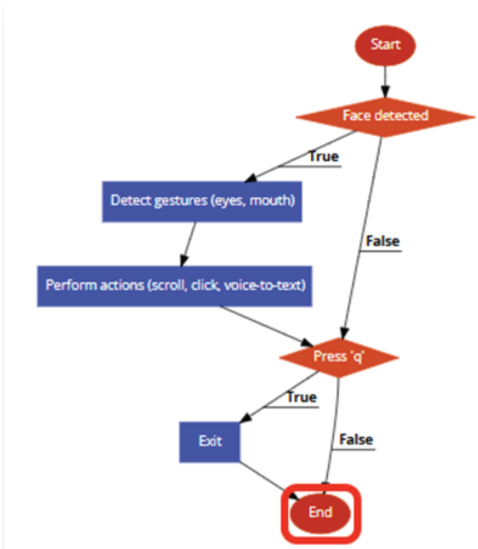

Figure 6: Flow chart

Usage

The following actions must be performed by the user to either enable or disable scrolling: closing both eyes for a certain time interval.

- Left Blink means' mouse click on left'.
- Right Blink means 'mouse click on right'.
- Scrolling up and down is achieved through head movement that is detected by vertical head movement of a person's nose tip position.
- Mouth Open activates the voice-to-text application.

Results and discussion

The test was conducted using the algorithm on the web page with anticipated outcomes. Through the camera, the face of the user was detected, indicating the most significant face regions in that area around the eyes and the mouth to represent input areas. Movement of the head would work as a mouse pointer-cursor in which a person moves their head left or right to move the cursor in the same direction. To activate scroll mode, the user closes both eyes momentarily to enable it, and head movement up or down would allow scrolling. A left blink causes theleft mouse click toperform, whilea right blink signifies a right mouse click; opening the mouth activates the voice-to-text feature, which allows converting spoken words into text.Indeed, the system recognizes.

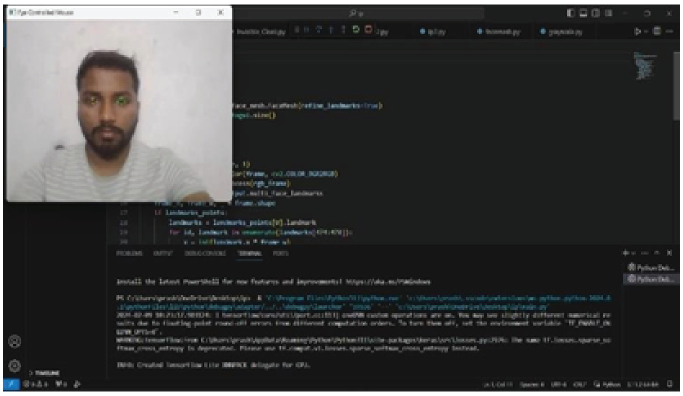

Figure 7: Output Image

References

1. Soukupová, Tereza, and Jan Cech. "Eye blink detection R. Bhallamudi et al., "Time and Statistical Complexity of Proposed Evolutionary Algorithm in Artificial Neural Networks," Pakistan Heart Journal, vol. 56, pp. 1014–1019, 2023.
2. R. Krishna et al., "Smart Governance in Public Agencies Using Big Data," The International Journal of Analytical and Experimental Modal Analysis (IJAEMA), vol. 7, pp. 1082–1095, 2020.
3. N. M. Krishna, "Object Detection and Tracking Using YOLO," in 3rd International Conference on Inventive Research in Computing Applications (ICIRCA-2021), IEEE, Sept. 2021, ISBN: 978-0-7381-4627-0.
4. N. M. Krishna, "Deep Learning Convolutional Neural Network (CNN) with Gaussian Mixture Model for Predicting Pancreatic Cancer," Springer US, vol. 1380-7501, pp. 1–15, Feb. 2019.
5. N. M. Krishna, "Emotion Recognition Using Skew

Gaussian Mixture Model for Brain–Computer Interaction," in SCDA-2018, Textbook Chapter, ISBN: 978-981-13-0514, pp. 297–305, Springer, 2018.

6. N. M. Krishna, "A Novel Approach for Effective Emotion Recognition Using Double Truncated Gaussian Mixture Model and EEG," I.J. Intelligent Systems and Applications, vol. 6, pp. 33–42, 2017.
7. N. M. Krishna, "Object Detection and Tracking Using YOLO," in 3rd International Conference on Inventive Research in Computing Applications (ICIRCA-2021), IEEE, Sept. 2021, ISBN: 978-0-7381-4627-0.
8. T. S. L. Prasad, K. B. Manikandan, and J. Vinoj, "Shielding NLP Systems: An In-depth Survey on Advanced AI Techniques for Adversarial Attack Detection in Cyber Security," in 2024 3rd International Conference on Automation, Computing and Renewable Systems (ICACRS), IEEE, 2024.
9. S. Sowjanya et al., "Bioacoustics Signal Authentication for E-Medical Records Using Blockchain," in 2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS), vol. 1, IEEE, 2024.
10. N. V. N. Sowjanya, G. Swetha, and T. S. L. Prasad, "AI Based Improved Vehicle Detection and Classification in Patterns Using Deep Learning," in Disruptive Technologies in Computing and Communication Systems: Proceedings of the 1st International Conference on Disruptive Technologies in Computing and Communication Systems, CRC Press, 2024.
11. C. V. P. Krishna and T. S. L. Prasad, "Weapon Detection Using Deep Learning," Journal of Optoelectronics Laser, vol. 41, no. 7, pp. 557–567, 2022.
12. T. S. L. Prasad et al., "Deep Learning Based Crowd Counting Using Image and Video," 2024.using facial landmarks." 21st computer vision winter workshop, RimskeToplice, Slovenia. 2016.
13. Chin, Craig A., and Armando Barreto. "Enhanced hybrid electromyogram/eye gaze tracking cursor control system for hands free computer interaction." 2006 InternationalConference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2006.
14. Jacko, Julie A., ed. Human computer interaction handbook: Fundamentals, evolving technologies, and emerging applications. CRC press, 2012.
15. Baltrušaitis, Tadas, Peter Robinson, and Louis-Philippe Morency. "Openface: an open source facial behavior analysis toolkit." 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2016.
16. Ballard, Philippe, and George C. Stockman. "Computer operation via face orientation." INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION. IEEE COMPUTER SOCIETY PRESS, 1992.
17. C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, M. Pantic. 300 Faces in the-Wild Challenge: The first facial landmark localization Challenge. Proceedings of IEEE Int'l Conf. on Computer Vision (ICCV-W), 300 Faces in-the-Wild Challenge (300-W). Sydney, Australia, December 2013.
18. Balamurugan, P., J. Santhosh, and G. Arulkumaran. "HAND MOTION BASED MOUSE CURSOR CONTROLUSING IMAGEPROCESSINGTECHNIQUE."Journalof
19. CriticalReviews7.4(2020):181-185.
20. Vasisht, Vinay S., Swaroop Joshi, and C. Gururaj. "Human computer interaction based eye controlled mouse." 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA). IEEE, 2019.
21. Mehta, Sukrit, et al. "Real-Time Driver Drowsiness Detection System Using Eye Aspect Ratio and Eye Closure Ratio." Available at SSRN 3356401 (2019).
22. Pimplaskar, D., Nagmode, M. S., & Borkar, A. (2015). Real time eye blinking detection and tracking using opencv. technology, 13(14), 15.
23. Kalas, M. S. (2014). Real Time Face Detection and Tracking Using OPENCV. international journal of soft computing and Artificial Intelligence, 2(1), 41-44.
24. Chau, M., & Betke, M. (2005). Real time eye tracking and blink detection with usb cameras. Boston University Computer Science Department.
25. R. Bhallamudi et al., "Deep Learning Model for Resolution Enhancement of Biomedical Images for Biometrics," in Generative Artificial Intelligence for Biomedical and Smart Health Informatics, Wiley Online Library, pp. 321–341, 2025.
26. R. Bhallamudi et al., "Artificial Intelligence Probabilities Scheme for Disease Prevention Data Set Construction in Intelligent Smart Healthcare Scenario," SLAS Technology, vol. 29, pp. 2472–6303, 2024, Elsevier.
27. R. Bhallamudi, "Improved Selection Method for Evolutionary Artificial Neural Network Design," Pakistan Heart Journal, vol. 56, pp. 985–992, 2023.