Proceedings of the International Conference on Sustainable Computing and Data Communication Systems (ICSCDS-2022)
IEEE Xplore Part Number: CFP22AZ5-ART; ISBN: 978-1-6654-7884-7

2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS) | 978-1-6654-7884-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICSCDS53736.2022.9760914

Mouse Cursor Control Using Facial Movements - An HCI Application

[1]Shashidhar R
Department of Electronics and Communication Engineering
Sri Jayachamarajendra College of Engineering, JSS Science & Technology University
Mysuru, INDIA
shashidhar.r@sjce.ac.in

[2]Snehith K
Department of Electronics and Communication Engineering
Sri Jayachamarajendra College of Engineering, JSS Science & Technology University
Mysuru, INDIA
snehithgowda27@gmail.com

[3]Abhishek P K
Department of Electronics and Communication Engineering
Sri Jayachamarajendra College of Engineering, JSS Science & Technology University
Mysuru, INDIA
20abhishekkamath@gmail.com

[4]Abhishek B Vishwagna
Department of Electronics and Communication Engineering
Sri Jayachamarajendra College of Engineering, JSS Science & Technology University
Mysuru, INDIA
abhivishwagna1@gmail.com

[5]Pavitha N
Department of Artificial Intelligence and Data Science
Vishwakarma Institute of Technology,
Pune, Maharashtra, INDIA
pavitha.n@vit.edu

***Abstract*— Because of various limitations, some people are unable to operate computers. Cursor controls based on facial expression are extremely useful not just for the future of natural input, but also for the handicapped and disadvantaged. Furthermore, creating a control system allows them to operate computers without the assistance of a third party. While this is not an issue for a healthy person, it may be an insurmountable barrier for persons who have a limited range of motion in their limbs. In these instances, input techniques based on the region's greater skills, such as eye movements and facial expressions, would be preferable. A system was created that uses a low-cost way to permit alternative input methods like controlling a mouse cursor on a computer screen. Using a camera we can capture the image of the eye and head movement. The mouse cursor is controlled by head movement as moving face up, down, sidewise movements and mouse events such as right and left click are controlled through eye blinks and squinting. This system is mainly aimed at disabled people to have effective communication with the computer. It is quite beneficial to implement a controlling system that allows them to move without the assistance of another person.**

Keywords— ***Computer-Vision, Face Landmark-recognition, Virtual mouse pointer.***

I. Introduction

With the boom in technology, Computers and their technologies have become important in the last few years for every aspect of our daily activities. Such behaviors vary from technical tasks to personal uses like surfing the internet, shopping, socializing, and entertainment. Computers are built for ordinary people. Computer input devices are typically inaccessible to those with significant physical limitations. It is difficult, if not impossible, for handicapped persons to utilize this technology effectively. Many persons who are unable to use their hands owing to a handicap or who are unable to speak to deliver voice commands are unable to evaluate computers, standard keyboards, and mouse. As a result, this community is deprived of numerous possibilities to connect and engage with others. For this reason, computer assistive tool design research is gaining traction and becoming increasingly important.

However, by combining both eyes tracking technology and directing cursor action with facial movement, this research attempts to build and implement an appropriate and robust human-computer interaction (HCI) interface for impaired users. We can control the cursor location on the computer screen by moving our eyes, head, and blinking. Following the user's input, real-time video input is analyzed, and facial recognition is conducted. The system is compatible with both desktops and laptops. This technology allows you to do things like move the pointer in all directions, scroll, drag, and click. As a result, the application would be more accurate and beneficial to physically handicapped users.

Real-time head tracking and eye state detection systems that employ a video camera face several difficulties. Light circumstances can vary, people might have varied face forms, and eyes can have different shapes and sizes depending on ethnicity when it comes to head mouse controls. As a result, the algorithm created proved to be resistant to various head orientations and lighting conditions. The goal of this research is to use image processing techniques from video recording to enhance the performance of recognizing head postures and eye states.
The main contribution of the proposed work is

- Develop the hand-free mouse controlling system.
- Develop the vision-based system.
- Facial gesture controls the mouse's activities.
- To get rid of the limitations of having an immobile head.

- Develop real-time eye tracking.

The rest of the article is as follows: Section II describes the merits and demerits of the existing works. Section III explains in detail the proposed methodology. Section IV discusses the output of the proposed model and comparison with existing work. Section V concluded the work with future scope.

II. Literature Survey

This paper's main goal is to assist persons who have difficulty utilizing their hands. The technology uses an eye-based interface to translate different eye movements into mouse cursor movements [1]. This system comprises a variety of software and libraries, including Python, OpenCV, and NumPy, as well as several distinct groups for face recognition. The Histogram of orchestrated Gradients is used to create the face identifier. This system does not rely on any external sensors or equipment. [2] The major goal of this paper is to analyze how to use a webcam to study various face gestures and movements on a personal computer. This is very important for physically handicapped individuals. The authors are employing a face detection algorithm to help in the estimation of the numerous parameters for high-precision computer control The cursor is controlled by tracking the position of the mouth, and page scrolling is accomplished by detecting the smile of the face.

The process of blink detection is discussed in this work[3]. This detection employs three kinds of eye blinks: left blink, right blink, and both. The blink detection procedure is separated into four steps: Facial images are captured via a web camera, followed by eye localization and motion analysis utilizing the optical flow technique. The authors of this work employed the MATLAB image processing and computer vision toolbox with MATLAB software. [7] The usage of a human-machine interface system for mouse control is proposed in this paper. The recognition system that has been proposed is based on a convolutional neural network, which uses low-resolution images recorded by a webcam [8] the user can move the cursor to the needed coordinates by moving his or her head in this system. The convolutional neural network converts head movements into actual cursor coordinates in the computer. This system is reliable and accurate to 92 percent.

This work describes a color filtering-based method for real-time face detection. Color filtering or motion filtering is employed in this system to examine each scale and location within the image. This method employs vision-based eye tracking, in which the movements and tracking of the eyes, as well as the gaze movements, are tracked. The system is used to determine visual attention and facilitate various user interaction tasks[9]. [13] The authors of this paper used an iris movement algorithm to control cursor movements in a computer. A facial recognition system will be used to determine the iris position, which will then be mapped on a computer screen. This makes it easier for physically challenged people to utilize computers[15]. The software is meant to process cursor moves up and down and also left and right with the help of the Iris movement[16]. Nowadays artificial Intelligence applying to all the applications like helmet detection, Vehicle number plate detection, and so on[17,18]. Vehicle number plate detection using machine learning and deep learning algorithm is a trend[19].

Hadish et al worked on wheelchair control using an eye-tracking system[20]. Kottilingam developed the robot using Multilanguage for elderly people[21]. Anushree proposed the deep neural network for the movement of the eye[22].

III. Methodology

A. Proposed HCL System using Facial Movements

The proposed architecture will allow users of various types to control the mouse cursor using facial movements. The system works from a normal webcam or the laptop's camera and it is hands-free where the user neither requires wearable hardware nor sensors. The actions include: Opening the mouth for activating HCL application, winking and squinting the eyes, moving the head around for upward and downward movements.

From Fig. 1, the image of the face is captured through the video stream which is taken from the webcam or video camera using the Face Detection Model. From the captured face, multiple modalities such as the Right eye, left eye, and Mouth is detected using the Landmark Detection Model. The head position movements which correspond to the movement of the mouse cursor are determined by using the Head Pose Estimation Model. Fig. 2 shows the proposed the actions and concerning functions are displayed.

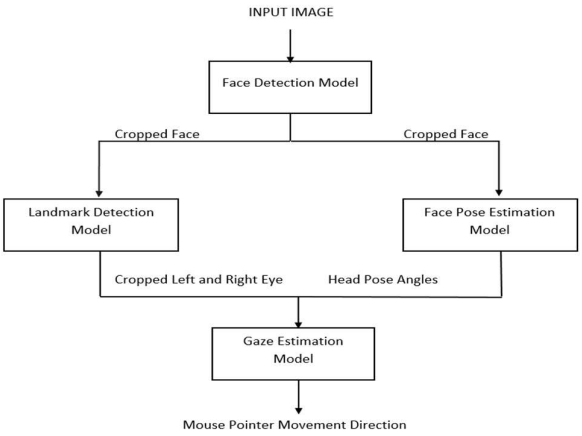

Fig. 1 Block Diagram of the Multimodal HCI System

The Gaze Estimation Model will be used to determine the direction of the mouse pointer by combining the outputs of both the Landmark Detection Model and the Head Pose Estimation Model. The Gaze Estimation model is in charge of facial motions for the mouse pointer to move properly.

Action	Function
 Opening Mouth	Activate / Deactivate Mouse Control
 Right Eye Wink	Right Click
 Left Eye Wink	Left Click
 Squinting Eyes	Activate / Deactivate Scrolling
 Head Movements (Pitch and Yaw)	Scrolling / Cursor Movement

Fig. 2 Proposed Set of Actions and Their Corresponding Functions.

B. *Proposed Algorithm for Multimodel HCL System*

The suggested model focuses on predicting a given face's facial landmarks. Many things are implemented using these landmarks from detecting eye blinks to predicting the emotions of the user. In this research work, a prebuilt model is used for fast detection of the face which allows the user to accurately predict 68 2D facial landmarks. The suggested approach is focused on predicting a given face's facial landmarks. Certain characteristics may be generated utilizing these expected landmarks of the face, allowing us to identify blinks using Eye-Aspect-Ratio and yawns using Mouth-Aspect-Ratio. To control the mouse cursor, these actions are written using the PyAutoGUI module.

C. *Eye Aspect Ratio[EAR]*

It helps in detecting blinks and winks of the eye. If the eye is closed, then the EAR value drops down and if the eye is open then the EAR value will be more than the set threshold. Eye aspect ratio as shown in fig. 3.
If EAR <=THRESHOLD:
STATUS = 'CLOSE'.

$$EAR = \frac{||p2 - p6|| + ||p3 - p5||}{2\,||p1 - p4||}$$

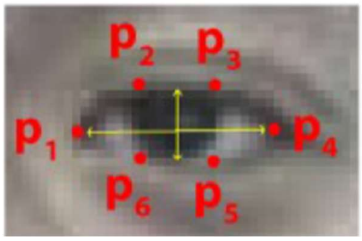

Fig. 3 Eye Aspect Ratio

D. *Mouth Aspects Ratio[MAR]*

It detects when a person's mouth opens and closes. When the mouth is open, the MAR value increases; otherwise, it drops. Multiple modalities are exploited and identified in the proposed system by incorporating both EAR and MAR. Mouth aspect ratio as shown in Fig. 4.
if MAR <=THRESHOLD:
STATUS = 'CLOSE'.

$$MAR = \frac{||p2 - p8|| + ||p3 - p7|| + ||p4 - p6||}{2\,||p1 - p5||}$$

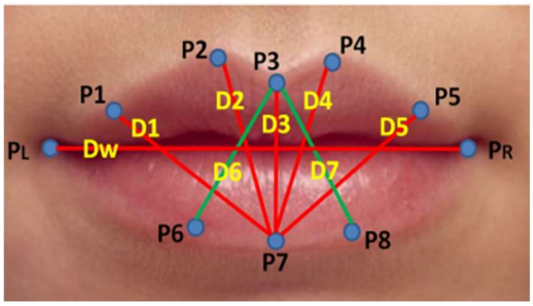

Fig. 4 Mouth Aspect Ratio

The model has two crucial features:
1. **Detector** - A device that detects the presence of a face.
2. **Predictor** - This model is used to predict landmarks.
Classic Histogram of Oriented Gradients (HOG) characteristics are combined with a linear classifier, sliding window detection, and an image pyramid to construct face detection. Detection of the features from the face using the proposed model as shown in Fig. 5.

Fig. 5 Face Features Detected Using Proposed Model.

The proposed algorithm for multimodal HCI system is specified below:
Algorithm:

1. Install the required libraries:
   a. Numpy.
   b. OpenCV.
   c. PyAutoGUI.
   d. Dlib.
   e. Imutils.
2. Initialize all the required thresholds, frame counters, and frame length for triggering mouse action.
3. Create a facial landmark predictor using the HOG-based Face Detector
4. Obtain the indexes of facial landmarks for eyes, nose, and mouth.
5. From the camera detect the face and convert it into grayscale channels.
6. Get facial landmark coordinates (x, y) and store them in the NumPy array.

7. Calculate Eye Aspect Ratio and Mouth Aspect Ratio.
8. Identify a nose point to draw a bounding box around it.
9. According to the drawn bounding box and actions, perform the required functions.

IV. RESULTS AND DISCUSSION

The experimental evaluation of the proposed system is carried out on Intel core i5 10th generation system with 8 GB RAM and 1 TB HDD on Windows 10 in PyCharm platform. The user can activate or deactivate the mouse control by opening the mouth. If the mouse control is active then READING INPUT! The message will be displayed on the system screen. The bounding box is drawn on the nose point by the system once mouse control gets activated. The direction of the mouse cursor can be varied by head movements (pitch and yaw). The mouse cursor movements in the left, right, up, and down directions are shown in Figures 6 and Figure 7.

The comparative analysis of the proposed system with existing systems is given in the table based on various parameters. All mouse control functions mentioned below are successfully implemented using the proposed algorithm of multimodal HCI system with facial movements. Table 1 discuss the comparative study of the proposed model with existing models.

1. Mouse cursor movement in left, right, up, and down directions.
2. Mouse scroll in up and down directions.
3. Mouse right-click and left-click.

TABLE 1: COMPARATIVE ANALYSIS OF PROPOSED SYSTEM WITH EXISTING SYSTEMS

System	Multimodal	Device Control Support	Supported Device Functions
Proposed Multimodal HCI	Yes	Yes	Mouse Cursor Movement (Left, Right, Up, Down). Mouse Scroll (Up, Down), Mouse Click (Right and Left)
Vision-Based Multimodal HCI	Yes	Yes	Mouse Cursor Movement (Left, Right, Up, Down). Mouse Click (Right and Left)
Eye-gaze Tracking	No	No	No
Eye Blink Monitoring	No	No	No

Our system aims to control the mouse motions and events hands-free by using face, eye blinks, and mouth. And our system can give the output as expected. Fig.6 shows the left and right movement of the cursor. Fig.7 shows the up and down movement of the cursor of the model and Fig.8 shows the up and down scrolling movements of the model. We got the result as follows:

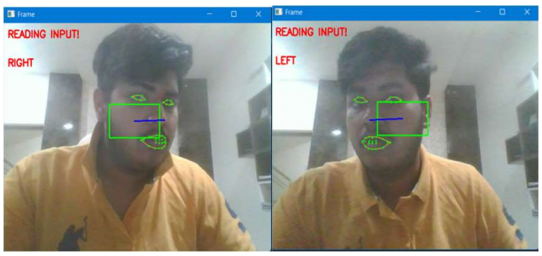

Fig. 6 Left and Right Movement of Cursor

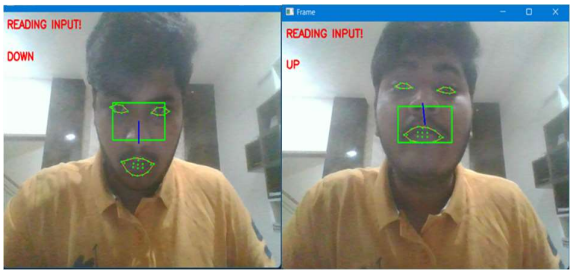

Fig. 7 Up and Down Movement of Cursor

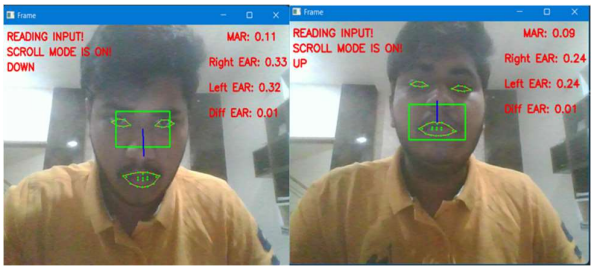

Fig. 8 Up and Down Scrolling Movements

V. CONCLUSION

In this paper, a Multimodal Human-Computer interface system had been designed and implemented. EAR and MAR features are used for parallel tracking of eyes and also for incorporating the multiple modalities by opening the respective mouth. The experimental evaluation of the proposed system is carried out on Intel core i5 10th generation system with 8 GB RAM and 1 TB HDD on Windows 10 in PyCharm platform. The proposed model works fine with the lower versions of specified computer features. All mouse control functions namely mouse cursor movement in left, right, up, and down directions; mouse scroll in up and down directions; and mouse right click and left click are successfully implemented using the proposed algorithm of multimodal HCI system. In the future, the present system can be extended to make the specified actions and their corresponding functions configurable. Other computer interfaced devices can also be controlled using HCI methodologies. The model can also be improved by considering only one facial movement when there are many faces in the frame.

REFERENCES

[1]. Bhagyashri P. Sonar and H. M. Baradkar, "Multimodal Human-Computer Interface System using Parallel Tracking of Eye" Journal of Engineering Technology Science and Innovation, Vol. 1, No. 1, April 2021.

[2]. Rohit Lal, Shital Chiddarwar, "Real-Time Human-Computer Interaction Using Facial Gestures" 10th ICCCNT 2019 July 6-8, 2019, IIT - Kanpur, Kanpur, India IEEE – 45670

[3]. Hari Singh, Jaswinder Singh, "Real-time eye blink and wink detection for object selection in HCI", © Springer International Publishing AG, part of Springer Nature 2018

[4]. Tereza Soukupova´ and Jan Cech, "*Real-Time Eye Blink Detection using Facial Landmarks"*. International 21st Computer Vision Winter Workshop, February 2018.

[5]. Wankhede, Shrunkhala Satish, S. Chhabria, and R.V. Dhar Askar. "Controlling mouse cursor using eye movement." International Journal of Application or Innovation in Engineering & Management, 2017.

[6]. Ing-Shiou Hwang, Yi-Ying Tsai, Bo-Han Zeng, Chien-Ming Lin, Huei-Sheng Shiue, Gwo-Ching Chang, "Integration of eye-tracking and lip motion for hands-free computer access" © Springer-Verlag GmbH Germany, part of Springer Nature 2020.

[7]. Rahib H. Abiyev, Murat Arslan, "Head mouse control system for people with disabilities", Special issue paper WILEY 2018

[8]. Akshada Dongre, Rodney Pinto, "Computer Cursor Control Using Eye and Face Gestures" 11th ICCCNT 2020 July 1-3, 2020 - IIT - Kharagpur, IEEE – 49239

[9]. G. Hu, Y. Xiao, Z. Cao, L. Meng, Z. Fang, and J. T. Zhou, "Towards real-time eyeblink detection in the wild: Dataset, theory, and practices," International Conference on Biomedical Engineering in Vietnam.

[10]. Bian, Z.P., Hou, J, Chau, L.P, Magnenat-Thalmann, N, "Facial position and expression-based human-computer interface for persons with tetraplegia", IEEE J. Biomed. Health Inform. 20(3), 915–924 (2016)

[11]. Jose, M., de Deus Lopes, R, "Human-computer interface controlled by the lip", IEEE J. Biomed. Health Inform. 19(1), 302–308 (2015)

[12]. Dey and Sanjay, "Real-Time Driver Fatigue Detection Based on Facial Behaviour along with Machine Learning Approaches". IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON), pp. 135-140, 2019.

[13]. K. Akhil Kumar, B R Ishwaryaa, Prasanna D V, Mrs. Shyamala. B, "VIRTUAL MOUSE USING EYE TRACKING", Journal of Xi'an University of Architecture & Technology, Volume XII, Issue IV, 2020.

[14]. Adrian Rosebrock "Eye blink detection with OpenCV, Python, and dlib," 2020.

[15]. Alhamzawi, H. A. (2018). Control mouse cursor by head movement: Development and implementation. Applied Medical Informatics Original Research, 40(3–4), 39–44.

[16]. Al-Rahayfeh, A., & Faezipour, M. (2013). Eye-tracking and head movement detection: A state-of-art survey. IEEE Journal of Translational Engineering in Health and Medicine, 1, 1–12.

[17]. Prajwal, M., et al. "A Review on helmet detection by using image processing and convolutional neural networks." *International Journal of Computer Applications* 182.50 (2019): 52-55.

[18]. MJ. Prajwal, KB. Tejas, V. Varshad, Murgod Mahesh Madivalappa, and R Shashidhar, "Detection of Non-Helmet Riders and Extraction of License Plate Number using Yolo v2 and OCR Method", International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 9, no. 2, December 2019.

[19]. R. Shashidhar, A. S. Manjunath, R. Santhosh Kumar, M. Roopa, and S. B. Puneeth, "Vehicle Number Plate Detection and Recognition using YOLO- V3 and OCR Method," 2021 IEEE International Conference on Mobile Networks and Wireless Communications (ICMNWC), 2021, pp. 1-5, doi: 10.1109/ICMNWC52512.2021.9688407.

[20]. Tesfamikael, Hadish Habte, Adam Fray, Israel Mengsteab, Adonay Semere, and Zebib Amanuel. "Simulation of Eye Tracking Control based Electric Wheelchair Construction by Image Segmentation Algorithm." Journal of Innovative Image Processing (JIIP) 3, no. 01 (2021): 21-35.

[21]. Kottilingam, Dr. "Emotional Wellbeing Assessment for Elderly Using Multi-Language Robot Interface." Journal of Information Technology and Digital World 2, no.1: 1-10.

[22]. Anusree K., Amudha J. (2020) Eye Movement Event Detection with Deep Neural Networks. In: Smys S., Tavares J., Balas V., Iliyasu A. (eds) Computational Vision and Bio-Inspired Computing. ICCVBIC 2019. Advances in Intelligent Systems and Computing, vol 1108. Springer, Cham. https://doi.org/10.1007/978-3-030-37218-7_98