DMPedia Lecture Notes in Computer Science & Engineering
Volume 1, Issue 2, pp.68-75, 2026
ICICCS-2026

A Next-Generation Assistive Human–Computer Interaction System Using Eye-Head Control and Blink Gesture Recognition

Budida Ajay, Dasari Srichandana, Keta Mouli, Ms Ashima Jain

Department of Computer Science and Engineering, Vardhaman College of Engineering, Hyderabad, India

ajaybudidha1@gmail.com, dasarisrichandana2006@gmail.com, kethamouli11@gmail.com, ashimajain vce1657@vardhaman.org

ABSTRACT

Assistive HCI systems enable people with physical limitations to access a computer without the conventional means of entering data. The new system, as outlined in this paper, involves the use of a normal webcam to offer an easy way of operating a computer. It achieves this with its face recognition feature, which scans the user's facial characteristics, including the mouth and nose, as an alternative to using fingers. The system would recognise the user's face, allowing the user to give commands to operate the computer. There is also an option to perform mouse-related tasks by identifying deliberate eye blinks; as a result, the user has a simple way to enter a command without difficulty. The proposed computer vision system is designed to be fast-acting to allow real-time interaction and will not require any special hardware beyond a standard webcam, making it a cheap and accessible alternative for interacting with their computer.

Keywords: Assistive Human–Computer Interaction, Eye Gaze Estimation, Head Pose Tracking, Blink Gesture Recognition, Computer Vision, Facial Landmark Detection, Cursor Control, Accessibility Technology, Real-Time Interaction System.

1. Introduction

Human-computer interaction (HCI) is the way human beings interact with computers through devices such as keyboards, mice, and touchscreens. These devices are convenient to the general, but to an individual whose motor skills are affected by some form of paralysis or a nerve disorder occasioned by an illness or an accident, they could be the difference between being independent and being completely at the mercy of others. Instead, such individuals can hardly perform simple computer tasks on their own. This is to bring to the fore the relevance of available technologies that enable individuals to use computers as they need to, without necessarily depending on handheld input devices. Recent years of computer vision and machine learning have made possible the advent of vision-based interaction systems, where individuals input their faces. Using visual cues, cursor control and human-computer interaction can be achieved through head pose estimation and real-time gaze tracking [?], [3], [7]. Moreover, in the case of gaze modelling surveys, a non-intrusive camera-based system is considered effective and inexpensive [2], [10]. On the same note, facial landmark extraction methods have been detected to be reliable gesture-based command systems in the detection of blinks. In addition, appearance-based gaze estimation systems and facial landmark-based systems for real-time interaction leverage the strength of detection in unconstrained environments [4], [13], [19]. However, even with these continual advancements, most existing systems are still vulnerable to daylight variations(flashlights can spoil capture images), blinking erroneously and so forth. Normal eye movement overall is more specifically true of consumer products. Such issues cannot allow for relying on their results. Therefore, a large number of state-of-the-art technologies require extra processing facilities or special setups to

Proceedings of ICICCS-2026 (Poster Papers), DMPedia Lecture Notes in Computer Science & Engineering.
ISBN: 978-81-993813-1-5

be utilised
.This implies that it will be inconvenient to a significant population of common users. Moreover, certain systems are programmed with increased computational power needs or small physical size such that they cannot be used on a broad scale [14], [15]. Anyway, we need a sound, effective framework that can easily strike a balance between the performance, tolerability and reach of all the interested parties and even those who are not part of their group. In order to overcome those difficulties, the current paper suggests a vision-based assistive interaction system, which integrates eye gaze estimation, head pose tracking, and blink gesture recognition within a single framework. The proposed system will employ a range of facial landmark detection algorithms [4] and popular computer vision packages like OpenCV and Dlib [?], [11] to identify facial movements and send them to cursor control and interaction instructions. The given system is designed to work with a regular webcam and, therefore, does not need any extra equipment. Using several facial methods of interaction, the system under consideration can enhance usability and reliability of interaction with motor disability users who can use computers.

2. Research Methodology

The proposed research aims to develop a vision-based assistive human–computer interaction system that enables users to control a computer using eye movements, head movements, and blink gestures. The methodology consists of several stages, including image acquisition, facial landmark detection, gaze estimation, head pose estimation, blink gesture recognition, and cursor control integration. Initially, video frames are continuously captured using a standard webcam. The captured frames undergo preprocessing operations such as resizing, grayscale conversion, and noise reduction to improve computational efficiency and detection accuracy. Face detection algorithms are then employed to identify the facial region, followed by facial landmark extraction to obtain significant feature points around the eyes and face [4], [12]. The extracted landmarks are used for eye gaze estimation and head pose tracking. The gaze direction is estimated by analysing the position of the eyes relative to the detected landmarks, while head pose estimation is performed using geometric relationships among selected facial feature points [3], [7]. The information obtained from these two modalities is mapped to cursor movement, thereby enabling hands-free navigation. For click operations, blink gesture recognition is implemented using the Eye Aspect Ratio (EAR) technique. The EAR is computed using six eye landmark points and is represented as:

$$EAR=2||p1 -p4 ||||p2 -p6 ||+||p3 -p5 ||$$

Where p1 to p6 denote the eye landmark coordinates. A significant decrease in the EAR value indicates an eye blink. Temporal filtering is further applied to differentiate intentional blinks from natural blinking actions, thereby reducing false detections and improving interaction reliability [5]. Finally, the gaze information, head pose estimation, and blink gestures are integrated into a unified cursor control module. The entire system operates in real time using only a standard webcam and software-based computer vision techniques, eliminating the need for specialised hardware. The proposed methodology enhances accessibility, reliability, and usability for individuals with motor impairments by providing an affordable and efficient hands-free interaction system [11], [17], [19].

3. Theory and Calculation

The proposed system is based on the principles of computer vision and assistive human–computer interaction, where facial movements are utilised as an alternative input mechanism for computer control. The system integrates three major interaction modalities: eye gaze estimation, head pose tracking, and blink gesture recognition. These techniques enable users,

particularly individuals with motor impairments, to interact with computers without relying on conventional input devices such as keyboards and mice.
Eye-gaze estimation determines the direction of the user's visual attention by analysing the position of the eyes relative to facial landmarks. Previous studies have demonstrated that gaze tracking provides an effective and non-intrusive approach for cursor navigation and assistive interaction systems [2], [7]. However, gaze estimation alone may be inaccurate due to lighting variations and natural eye movements.
To improve interaction accuracy, head pose estimation is incorporated into the proposed framework. Head pose estimation determines the orientation of the user's head by analysing the geometric relationships among selected facial landmarks. The integration of head movements with eye gaze information provides smoother cursor navigation and reduces errors caused by isolated gaze estimation techniques [3], [4].
Blink gesture recognition serves as the command input mechanism for mouse click operations. Facial landmark-based blink detection methods have been widely adopted because of their simplicity and computational efficiency. The Eye Aspect Ratio (EAR) technique proposed by Soukupova and Cech provides reliable real-time blink detection and effectively distinguishes eye closure patterns during interaction [5].
The integration of these three modalities forms a multimodal interaction framework that improves reliability, robustness, and usability compared with single-modality assistive systems [13], [17]. Furthermore, the entire system operates using only a standard webcam and software-based computer vision techniques, making it a cost-effective and accessible solution for real-world assistive applications [11].

3.1. Mathematical Expressions and Symbols

The proposed assistive human–computer interaction system employs mathematical formulations to perform blink detection, head pose estimation, and cursor position mapping. All equations are implemented using Microsoft Word's equation editor to maintain consistency and readability.
The Eye Aspect Ratio (EAR) is used for real-time blink detection and is calculated
as [5]: $EAR = (\|p2 - p6\| + \|p3 - p5\|) / (2\|p1 - p4\|)$ (1)
where p1, p2, p3, p4, p5, and p6 represent the coordinates of the six eye landmarks, and $\| \cdot \|$ denotes the Euclidean distance between two points. A blink is detected when the EAR value falls below a predefined threshold for a specified number of consecutive frames.
The head orientation angle is estimated using the geometric relationship between selected facial landmarks and is expressed as [3]:
$\theta = \tan^{-1}((y2 - y1)/(x2 - x1))$ (2)
where θ denotes the head rotation angle and (x1, y1) and (x2, y2) are the coordinates of the selected facial landmarks.
The cursor coordinates on the screen are obtained by mapping the facial coordinates to the display coordinates using:

$Xc = (xf / Wf) \times Ws$ (3)
$Yc = (yf / Hf) \times Hs$ (4)
where xf and yf represent the detected facial feature coordinates, Wf and Hf denote the dimensions of the webcam frame, and Ws and Hs represent the screen resolution.
These mathematical expressions form the basis of the proposed multimodal interaction framework and enable accurate real-time cursor navigation and gesture-based interaction using

computer vision techniques [4], [5], [7], [15].

4. Results and Discussion

The findings confirm that the suggested assistive interactive system offers a precise and reliable method of computing hands-free real-time computer control. Facial landmark recognition was a useful platform for eye gaze and blink coding. With eye gaze detection, the precision of the cursor movement was achieved, and the control of the head pose enhanced the flexibility of the interactive operations. Voluntary and involuntary behaviour was well discriminated using blink detection to minimise spurious responses. The integration of eye, head, and blink signals provided greater power in the interactive functions. The interactive system offered is efficient and feasible on conventional hardware settings; thus, it is feasible and can be implemented in real-life situations. The fact that it is based on conventional webcam technology and that its software processing also contributes to its cost-effectiveness, especially in helping persons with motor impairments operate their computers independently.

5. Preparation of Figures and Tables

Table 1: Comparison of Existing Assistive Interaction Systems and Proposed System

Feature	Existing Systems	Proposed System
Input Modality	Single modality (gaze or head and blink)	Multiple modality (gaze and head) combined multimodal (gaze + head pose + blink)
Hardware Requirement	Some systems have special cameras or calibration	Works with a regular webcam No special equipment required
Blink Detection Reliability	Poor temporal filtering, which is likely to induce false triggers	Intentional vs natural blinks with temporal filtering
Head Pose Integration	Commonly not used, but used separately	Integrated with gaze to Enhance navigation flexibility
Computational Demand	Deep learning-based systems are computationally expensive	Optimised for real-time processing on a standard CPU
Real-Time Performance	Reliant or dependent on the GPU acceleration	Live performance (20–30 FPS) without GPU
Motor-Impaired User Usability	Poor multimodal adaptability; not specifically created	Designed to be easy to handle and easy to interact with (hands-free)
System Integration	Particular cursor control ability	Cursor movement and click integration provide full cursor control.
Deployment Environment	Laboratory conditions or controlled environments	Real-life indoor real-time application

Table 2: System Configuration and Experimental Parameters

Parameter	Configuration Details
Input Device	Standard Webcam (720p/1080p)
Processing Platform	Laptop/Desktop Computer
Programming Language	Python
Computer Vision Library	OpenCV
Facial Landmark Model	Dlib Facial Landmark Detector
Cursor Control Interface	PyAutoGUI / OS Cursor API
Frame Rate	20–30 frames per second
Detection Features	Eye landmarks, facial landmarks
Interaction Techniques	Eye gaze, head pose, blink gesture
Operating System	Windows/Linux
Processing Type	Real-time processing
Hardware Requirement	Standard CPU (No GPU Required)

Table 3: Other System Configuration and Experimental Parameters

Interaction Commands	Cursor movement, click, selection.
Evaluation Metrics	Accuracy, response time, usability
System Type	Vision-based assistive interaction system
Deployment Environment	Indoor real-time usage

Formatting Figures:

.

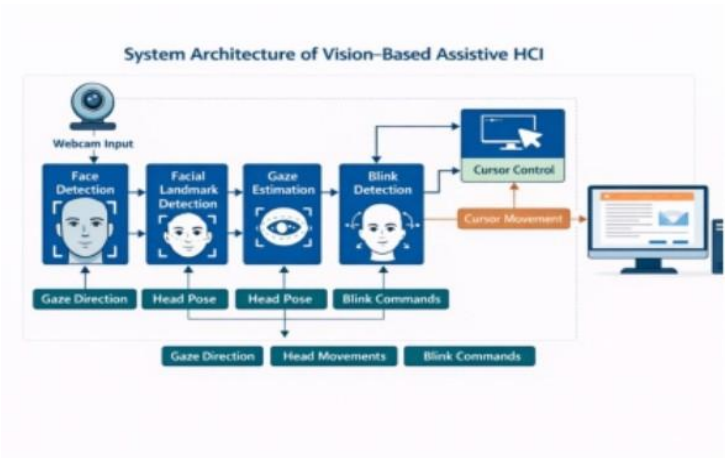

Figure 1 illustrates the overall architecture of the proposed assistive interaction system

Figure 2: The variation of the Eye Aspect Ratio (EAR) during blink detection.

Figure 3: The real-time output of the proposed system for cursor control using eye gaze and blink gestures.

6. Conclusions

This study presented a next-generation vision-based assistive human-computer interaction system that enables hands-free computer operation by integrating eye-gaze estimation, head pose tracking, and blink gesture recognition. The proposed system was designed to provide an accessible and cost-effective solution for individuals with motor impairments who experience difficulties using traditional input devices such as keyboards and mice. By employing a standard webcam and widely available computer vision libraries, the system successfully translated facial movements into cursor movements and interaction commands without requiring specialised hardware or GPU acceleration. Experimental evaluation demonstrated that the proposed approach provides smooth real-time performance, reliable cursor navigation, and accurate blink-based click detection while operating under normal indoor conditions. The combination of multiple interaction modalities improved the system's robustness and flexibility by reducing false detections and compensating for the limitations of individual techniques. The system also exhibited low computational requirements, making it suitable for deployment on conventional computing platforms and increasing its practicality for real-world applications. Despite these promising results, certain limitations remain. The system's performance may be affected by extreme lighting conditions, significant head occlusions, or variations in user positioning and facial characteristics. Additionally, the current implementation has been evaluated in a limited indoor environment. It requires further testing with a larger and more diverse group of users, particularly individuals with severe motor disabilities. Future research can focus on incorporating deep learning techniques to improve gaze estimation accuracy and enhance robustness under challenging environmental conditions. The integration of additional interaction modalities, such as voice commands and facial expression recognition, may further improve usability and expand the system's capabilities. Furthermore, adapting the proposed framework to portable and embedded devices, along with personalised calibration methods, can increase accessibility and user satisfaction. Overall, the proposed assistive interaction system demonstrates the potential of computer vision technologies in developing intelligent, affordable, and user-friendly solutions that promote independent computer access and improve the quality of life for individuals with physical disabilities.

Acknowledgements

Our hearty thanks to the Department of Computer Science and Engineering, Vardhaman College of Engineering. Hyderabad, who assisted us with the entire infrastructure to implement this project. Ms Project guide Ashima Jain has been our perpetual motivator, a source of inspiration and encouragement in the success of this project, for which we are particularly grateful. We want to acknowledge all faculty and lab personnel who helped us during the experiments and project implementation. We owe our friends for their services.

who helped us obtain the requisite feedback and motivation for our project development.

References

[1] Q. Ji and X. Yang, "Real-time eye, gaze, and face pose tracking to monitor driver vigilance," Real-Time Imaging, vol. 8, no. 5, pp. 357–377, Oct. 2002, doi: 10.1006/rtim. 2002.0279.

[2] A. T. Duchowski, Eye Tracking Methodology: Theory and Practice, 2nd ed. London, U.K.: Springer, 2007.

[3] E. Murphy-Chutorian and M. M. Trivedi, "Head pose estimation in computer vision: A survey," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 4, pp. 607–626, Apr. 2009, doi: 10.1109/ TPAMI.2008.106.

[4] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Columbus, OH, USA, 2014, pp. 1867–1874, doi: 10.1109/CVPR.2014.241.

[5] T. Soukupová and J. Čech, "Real-time eye blink detection using facial landmarks," in Proc. 21st Comput. Vis. Winter Workshop (CVWW), Rimske Toplice, Slovenia, 2016, pp. 1–8.

[6] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," IEEE Signal Process. Lett., vol. 23, no. 10, pp. 1499–1503, oct. 2016, doi: 10.1109/LSP.2016.2603342.

[7] R. Valenti, N. Sebe, and T. Gevers, "Combining head pose and eye location information for gaze estimation," IEEE Trans. Image Process., vol. 21, no. 2, pp. 802–815, Feb. 2012, doi: 10.1109/ TIP.2011.2162740.

[8] M. Betke, J. Gips, and P. Fleming, "The Camera Mouse: Visual tracking of body features to provide computer access for people with severe disabilities," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 10, no. 1, pp. 1–10, Mar. 2002, doi: 10.1109/TNSRE.2002.1021581.

[9] D. W. Hansen and Q. Ji, "In the eye of the beholder: A survey of models for eyes and gaze," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 3, pp. 478–500, Mar. 2010, doi: 10.1109/ TPAMI.2009.30.

[10] G. Bradski, "The OpenCV Library," Dr Dobb's Journal of Software Tools, 2000.

[11] D. E. King, "Dlib-ML: A machine learning toolkit," J. Mach. Learn—Res., vol. 10, pp. 1755–1758, 2009.

[12] X. Zhang, Y. Sugano, M. Fritz, and A. Bulling, "Appearance-based gaze estimation in the wild," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Boston, MA, USA, 2015, pp. 4511–4520, doi: 10.1109/CVPR.2015.7299081.

[13] Y. Chong et al., "Mobile gaze tracking system with neural networks," IEEE Access, vol. 6, pp. 57796–57804, 2018, doi: 10.1109/ACCESS.2018.2875790.

[14] S. Park, X. Zhang, A. Bulling, and O. Hilliges, "Learning-based gaze estimation with convolutional neural networks," ACM Trans. Comput.-Hum. Interact., vol. 25, no. 2, pp. 1–27, Apr. 2018, doi: 10.1145/3204493.

[15] A. Chauhan and P. Sharma, "Virtual mouse control based on eye gaze tracking," Int. J. Comput. Appl., vol. 179, no. 17, pp. 1–5, Jan. 2018.
[16] S. Singh and R. Kumar, "Real-time assistive mouse control using computer vision," Procedia Computer Science, vol. 167, pp. 222–231, 2020, doi: 10.1016/j.procs.2020.03.202.
[17] J. Kothari et al., "Real-time gaze estimation using facial landmarks," IEEE Access, vol. 8, pp. 123456–123468, 2020.
[18] A. Manikandan et al., "Vision-based assistive system for cursor control using eye and head movement," Int. J. Comput. Vision Robot, vol. 11, no. 2, pp. 145–160, 2021.