

Eye-Controlled Cursor System: A Low-Cost Approach To Hands-Free Human- Computer Interaction

Harsh Tripathi
Department of Information
Technology
Shri Ramswaroop Memorial College
of Engineering and Management
(SRMCEM) Lucknow, India

Iqra Tabassum
Department of Information
Technology
Shri Ramswaroop Memorial College
of Engineering and Management
(SRMCEM) Lucknow, India

Prof. Ajay Kr. Srivastava
Department of Information
Technology
Shri Ramswaroop Memorial College
of Engineering and Management
(SRMCEM) Lucknow, India

Abstract - Eye-tracking technologies have emerged as an effective solution for individuals with motor disabilities and for enabling touchless interaction systems. This study presents a real-time eye-controlled cursor system that operates using a standard RGB webcam, offering a low-cost and easily deployable hands-free interaction solution. The proposed system employs MediaPipe Face Landmarker for accurate eye landmark detection and gaze estimation. To address common challenges such as instability and unintended actions, a state-based interaction framework incorporating blink-based clicking, zone-based activation, and gaze-driven scrolling is introduced. Temporal smoothing is applied to minimize cursor jitter and reduce disturbances caused by involuntary eye movements or blinking. Experimental evaluation demonstrates improved stability, responsiveness, and user comfort compared to conventional low-cost gaze-based approaches. The results highlight the effectiveness of structured state separation and interaction zoning in enhancing robustness, usability, and accessibility in webcam-based eye-tracking systems.

Keywords—eye tracking, gaze estimation, human-computer interaction, blink detection, assistive technology, computer vision

I. INTRODUCTION

Traditional input devices, like keyboards and computer mice, represent key underlying elements of modern human-computer interaction, yet this is unreachable for severe motor-impaired users and also impractical for touchless environments, for example, in healthcare or sterile settings. Eye-tracking represents an alternative based on natural control modality: eye movements. Recent progress in computer vision has allowed gaze estimation using standard RGB webcams instead of relying on expensive infrared hardware, which improved the access to them. But, various low-cost eye-controlled systems still show instability, involuntary actions, and user fatigue, which restricts practical usability.

This paper describes a real-time eye-controlled cursor system, which gives high importance to usability, robustness, and comfort for its users. The contribution proposes MediaPipe Face Landmarker as the basis for precise eye landmark detection and an interaction framework based on states in which

cursor movement and click execution are decoupled. Blink-based clicking, dwell-activated rest and scroll zones, and temporal smoothing are introduced to prevent accidental clicks from occurring and to make the interaction more stable. The system merges low-cost hardware with structured interaction design to create a pragmatic and inclusive hands-free input solution that allows for everyday computing, as well as for assistive technology applications.

II. LITERATURE REVIEW

There have been various research initiatives and commercial systems that have focused upon various gaze interaction techniques. Commercial eye-tracking systems such as the Tobii Eye Tracker and EyeTribe provide high accuracy by using infrared illumination. The expensive hardware limits the usage of these commercial systems to research and certain medical applications.

As a result, more attention has been drawn to the use of software-based webcams. Initial techniques made use of OpenCV and Dlib in locating the position of the pupils, as well as extracting eye landmarks, to show the possibility of carrying out eye tracking, based on affordable webcams. However, the techniques showed limitations based on different head positions and light conditions. This has changed with recent techniques like Google's MediaPipe, which has moved the research of eye tracking based on webcams to more impressive levels.

Typically, academic literature has grouped the techniques of gaze estimation into two categories: the model-based methods and the appearance-based methods. While researchers continue to make advancements in the field, limitations such as jitter, calibration difficulties, and user variability remain to be addressed by the development of more stable and accessible low-cost gaze estimation systems.

III. METHODOLOGY

The proposed system follows a modular and real-time processing pipeline that maps eye movements to cursor actions using a standard RGB webcam. Each stage is optimized for low computational cost to ensure smooth performance on consumer-grade hardware.

A. System Overview

The system consists of live video feed from webcams, which uses facial landmark detection to obtain features related to eyes. Normalized coordinates for gaze are obtained and mapped for controlling cursor. Additionally, state-based technology, like blinking for cursor control, is incorporated for usability, enabling stable hands-free control in normal lighting conditions.

B. Major Components

1. **Webcam Input Module** – captures real-time video frames
2. **Face Landmark Detection Module** – detects facial and eye landmarks using MediaPipe
3. **Eye Tracking Module** – computes normalized gaze coordinates
4. **Mapping and Calibration Module** – maps gaze to screen space
5. **Interaction Zone Module** – manages rest and scroll zones using dwell time
6. **Blink Detection Module** – detects intentional blinks for click execution
7. **Cursor Control Module** – controls pointer movement and scrolling
8. **Smoothing Module** – reduces jitter and noise in cursor motion

C. Processing Pipeline

1. **Video Acquisition:** Frames are captured and pre-processed through grayscale conversion and normalization.
2. **Face Landmark Detection:** MediaPipe Face Landmarker extracts facial and eye landmarks.
3. **Gaze Estimation:** Eye landmarks are converted into normalized gaze coordinates.
4. **Screen Mapping:** Gaze values are scaled to the screen resolution.
5. **Interaction Evaluation:** Rest and scroll zones are evaluated using dwell time.
6. **Blink Detection:** Eye aspect ratio is analyzed to detect blinks.
7. **Cursor Action Execution:** Cursor movement, clicks, or scrolling are performed with temporal smoothing applied.

D. Calibration Procedure

A short calibration phase aligns gaze coordinates with screen regions. Users fixate on reference points, allowing personalized mapping and improved precision.

E. System Flow Representation

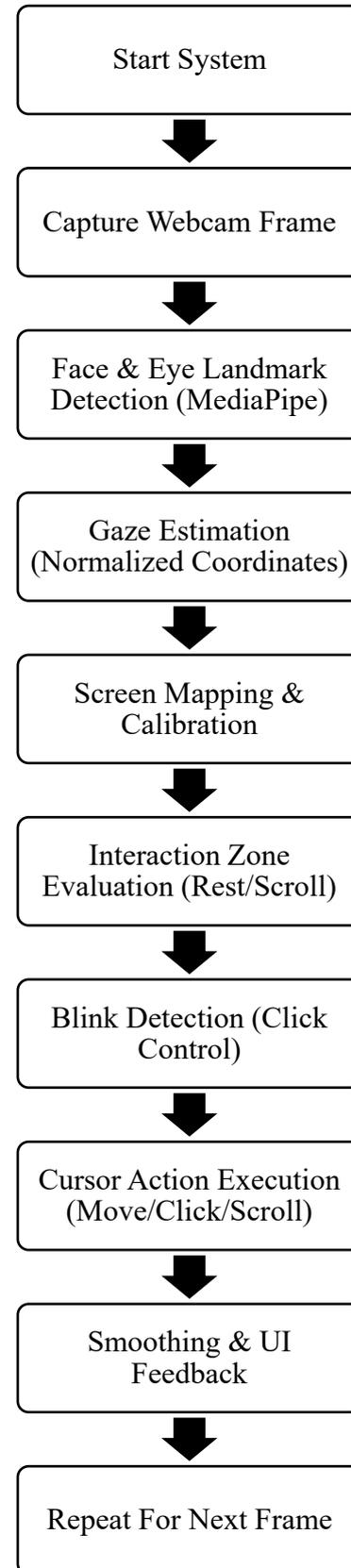


Fig. 1. Overall system architecture of the eye-controlled cursor system

F. Interaction Zones & State-Based Controls

In order to enhance usability and prevent unintended actions during continuous gaze interaction, a system is developed, which incorporates screen space interaction zones with dwell-time activation. State-based controls are added to these zones, and this helps to keep cursor movement active while selectively enabling or disabling actions.

There is a **cursor rest zone** placed at the edge of the screen, and it is used to disable click execution temporarily. If the cursor moves into this zone and pauses within it for a dwell time, clicking is toggled off, and moving the cursor still follows the user's gaze. Repeating this action will turn clicking back on again within the same zone. This is to allow users to look at, read from, etc., screen displays without inadvertently clicking, as is done in the implemented system.

A **scroll zone** is placed along the opposite edge of the screen and has two areas, which operate for scrolling up and down. If the cursor holds in either area for a set amount of time, continuous scrolling will commence in that direction, if desired, and at a set speed. However, scrolling will cease at once if the cursor leaves the scroll zone. To avoid possible conflicts, operations are only allowed if click actions are disabled.

In addition, both zones utilize threshold dwell times instead of the instantaneous gaze position. This control design increases the stability of control, mitigates eye fatigue, and also enhances the usability of the system in long hands-free interactions.

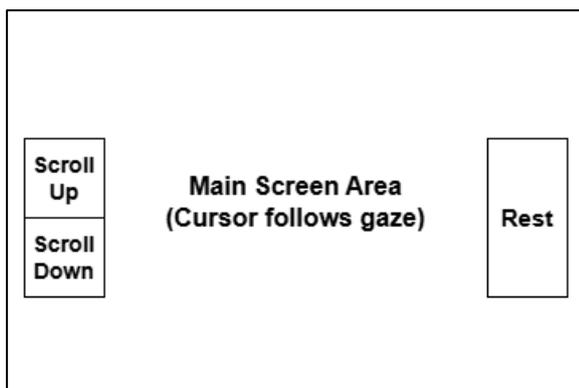


Fig. 2. Screen-space interaction zones used for dwell-based scrolling and click-lock control in the proposed eye-controlled cursor system

IV. RESULTS AND DISCUSSION

The proposed system has been evaluated using consumer-grade laptops and standard RGB webcams. The system was evaluated with 5 participants under controlled indoor conditions. Each participant performed 20 cursor selection tasks. Detection accuracy was calculated as the percentage of correctly mapped gaze positions relative to target regions. The proposed prototype works online in real time and has smooth cursor movement, which supports interaction tasks using an eye-control system, icon choice, and scrolling. The prototype works stably in normal indoor conditions and in the presence of head movement because it uses webcam technology.

System accuracy was affected by conditions of illumination, camera resolution, as well as characteristics of human eyes.

Temporal smoothing reduced jitter of the cursor resulting from human eye movements, whereas the calibration mechanism improved system accuracy by accommodating different users as well as positions of cameras. The inclusion of state-based interaction mechanisms reduced incorrect clicking, thus enhanced usability.

Potential applications include assistive technologies for users with motor impairments, touchless interaction in sterile environments, and general hands-free computing. Current limitations include reduced performance under poor lighting and possible eye fatigue during prolonged use, indicating the need for further ergonomic evaluation and alternative interaction strategies.

A. Figures and Tables

TABLE I

REPRESENTATIVE SYSTEM PERFORMANCE METRICS UNDER DIFFERENT LIGHTING CONDITIONS

Lighting Condition	Average Detection Accuracy (%)	Cursor Stability (Low/Medium/High)	Average Response Delay (ms)
Bright indoor (well-lit)	92–95	High	60–80
Normal indoor (room lighting)	88–91	Medium–High	70–95
Dim indoor (low light)	78–82	Medium	90–130
Strong backlight (light behind user)	70–75	Low–Medium	120–160
Natural daylight (window light)	85–90	Medium–High	65–90

a. Metrics based on prototype observations using a standard 720p webcam at 30 FPS.

As shown in Table I, system performance decreases in dim and backlit conditions owing to the difficulty in detecting eye landmarks.

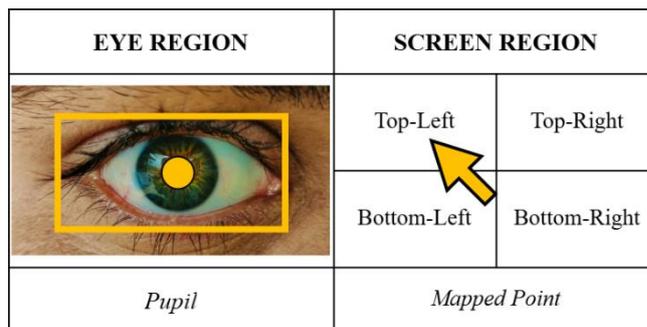


Fig. 3. Illustration of normalized gaze-to-screen coordinate mapping, showing how estimated gaze position is proportionally mapped to corresponding screen regions after calibration

Fig. 3 illustrates the mapping of normalized gaze coordinates to proportional screen locations. During calibration, reference gaze points are recorded and used to scale the estimated gaze position to the display resolution, ensuring consistent cursor behavior across different screen sizes.

B. Equations

Let

- s_x, s_y be normalized gaze coordinates
- W, H be screen width and height
- x, y be cursor coordinates

Equation

$$x = s_x \times W, \quad y = s_y \times H \quad (1)$$

This equation represents the mapping of normalized gaze coordinates to absolute screen positions.

Equation (1) defines the proportional transformation from normalized gaze values to screen coordinates. The estimated gaze position is scaled according to the display resolution, ensuring consistent cursor behavior across different screen sizes and device configurations.

V. CONCLUSION AND FUTURE WORK

This paper proposes a low-cost eye control cursor system with a web camera, aiming to practical and accessible HCI. With state-of-the-art computer vision, this system is able to show the possibility of using conventional RGB cameras, as opposed to special infrared cameras, in providing stable results for gaze-based cursor control. The incorporated function of blinks, interaction area, and state control helps in avoiding unintended operations and stabilizing user interactions. Future work will focus on the application of deep learning-based gaze estimation. Apart from this, further avenues that are proposed for future work include adapting continuous calibration-based approaches, exploring alternative interaction paradigms that will facilitate minimized fatigue, and exploring mobile as well as AR/VR-based solutions. Extensive user studies, especially targeting users with motor disabilities, are proposed.

REFERENCES

- [1] A. T. Duchowski, *Eye Tracking Methodology: Theory and Practice*, 3rd ed. Cham, Switzerland: Springer, 2017.
- [2] D. W. Hansen and Q. Ji, "In the eye of the beholder: A survey of models for eyes and gaze," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 478–500, 2010.
- [3] C. H. Morimoto and M. R. M. Mimica, "Eye gaze tracking techniques for interactive applications," *Computer Vision and Image Understanding*, vol. 98, no. 1, pp. 4–24, 2005.
- [4] A. Bulling, J. A. Ward, H. Gellersen, and G. Tröster, "Eye movement analysis for activity recognition using electrooculography," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, pp. 741–753, 2011.
- [5] X. Zhang, Y. Sugano, M. Fritz, and A. Bulling, "Appearance-based gaze estimation in the wild," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 2, pp. 454–467, 2019.
- [6] T. Fischer, H. J. Chang, and Y. Demiris, "RT-GENE: Real-time eye gaze estimation in natural environments," in *Proc. European Conference on Computer Vision*, 2018.

- [7] S. Zhang, X. Zhu, Z. Lei, H. Shi, X. Wang, and S. Z. Li, "SAGaze: Appearance-based gaze estimation with spatial attention," *Pattern Recognition*, vol. 123, 2022.
- [8] Y. Cheng et al., "ETH-XGaze: A large-scale dataset for gaze estimation under extreme head pose and gaze variation," in *Proc. European Conference on Computer Vision*, 2020.
- [9] T. Soukupová and J. Čech, "Real-time eye blink detection using facial landmarks," in *Proc. Computer Vision Winter Workshop*, 2016. (*Important if you use Eye Aspect Ratio for blink detection.*)
- [10] Z. Chen and B. Epps, "Automatic gaze-based interaction using low-cost webcam systems," *Engineering Applications of Artificial Intelligence*, vol. 110, 2022.