AI Eye & Hand Gesture Control System – Complete Project Explanation

This project was developed for the RIFT Hackathon. It demonstrates a complete touch-free human–computer interaction system using computer vision. The system allows users to control mouse movement, clicking, scrolling, and screen brightness using eye movements and hand gestures through a single webcam.

1. Project Overview

The project uses real-time face and hand tracking to interpret human gestures as computer input. Eye movement controls the mouse pointer, blinking performs mouse clicks, hand gestures manage scrolling, freezing actions, and brightness control. The system is designed with accessibility and assistive technology in mind.

2. Tools and Technologies Used

• Python – Core programming language
• OpenCV – Webcam access and frame processing
• MediaPipe Face Mesh – Eye and face landmark detection
• MediaPipe Hands – Hand gesture and finger tracking
• PyAutoGUI – Mouse and scroll automation
• screen-brightness-control – System brightness adjustment
• NumPy – Mathematical calculations and interpolation

3. System Workflow

The webcam captures live video frames. MediaPipe processes each frame to detect face and hand landmarks. Eye landmarks are used to control cursor movement and blink detection. Hand landmarks are analyzed for gesture recognition such as fist detection and finger pinch. Based on these gestures, corresponding system actions are executed in real time.

4. Camera Initialization

OpenCV is used to access the webcam using VideoCapture. Frames are continuously read inside a loop, flipped for mirror view, and converted to RGB format before being passed to MediaPipe models.

5. Eye Tracking and Mouse Movement

The iris landmark coordinates are mapped to the screen resolution. These mapped values are used to move the mouse cursor smoothly across the screen using PyAutoGUI.

6. Blink Detection for Mouse Click

Blinking is detected by measuring the distance between upper and lower eyelid landmarks. When the distance falls below a threshold, it is considered a blink and triggers a mouse click. A cooldown mechanism prevents accidental multiple clicks.

7. Freeze Mode Using Fist Gesture

When all fingers are folded into a fist, the system toggles freeze mode. In freeze mode, mouse movement stops, allowing the user to perform scrolling actions without moving the cursor.

8. Scrolling Using Eye or Head Movement

While freeze mode is active, vertical eye or head movement is mapped to scroll up or down actions. This allows hands-free scrolling of documents or web pages.

9. Brightness Control Using Finger Pinch

The distance between the thumb tip and index finger tip is calculated. This distance is mapped to a brightness range of 10% to 100% using interpolation. The brightness is updated dynamically in real time.

10. Displaying Brightness Percentage

The current brightness percentage is displayed on the camera window. This provides real-time visual feedback to the user while adjusting brightness.

11. User Interface Feedback

The camera window displays the current system mode (NORMAL or FREEZE), brightness percentage, and visual markers for detected fingers. This improves usability and clarity during interaction.

12. Exit and Resource Handling

The application runs continuously until the user presses the 'Q' key. All camera resources and windows are released properly to ensure a clean exit.

13. Project Outcome

The project successfully demonstrates a complete hands-free and touch-free computer control system. It performs reliably in real time and can be used on standard laptops with a single webcam.

14. Applications

• Assistive technology for people with motor disabilities
• Touchless computer interaction
• Smart workspaces and AR/VR interfaces
• Human–Computer Interaction (HCI) research
• Hackathon and academic demonstrations

15. Conclusion

This project showcases the power of computer vision in building intuitive and accessible systems. By combining MediaPipe, OpenCV, and Python automation libraries, complex human gestures can be translated into meaningful digital actions. The system is scalable, extendable, and suitable for real-world applications.