Real-Time Hand Gesture Recognition using MediaPipe and OpenCV for Virtual Human-Computer Interaction

Amaravathi Veerendra Nath, Sri Varsha Tandra

Abstract

This paper presents a real-time hand gesture recognition system developed using MediaPipe and OpenCV frameworks to enhance virtual human-computer interaction. The proposed system detects hand landmarks and tracks finger positions to interpret gestures efficiently. Our implementation improves upon existing solutions by optimizing detection accuracy and computational efficiency, making it suitable for real-time applications.

Keywords

Hand Gesture Recognition, MediaPipe, OpenCV, Computer Vision, Human-Computer Interaction

I. Introduction

Hand gesture recognition is an important area in human-computer interaction, enabling intuitive control without physical contact. Recent advancements in computer vision and machine learning have facilitated the development of robust hand tracking systems. MediaPipe, a framework by Google, provides efficient hand landmark detection, which, when combined with OpenCV, can be leveraged for real-time gesture recognition. This paper details the development of such a system, highlighting the methodology, implementation, and enhancements introduced.

II. Methodology

The system is built using Python and primarily utilizes MediaPipe's Hands module for detecting and tracking 21 hand landmarks in real-time video streams. OpenCV is employed for image processing, visualization, and capturing video feed.

The handDetector class encapsulates the detection logic, providing functionalities to find hand landmarks, determine finger states, and calculate distances between landmarks for gesture recognition.

Key features include:

- Real-time detection with adjustable confidence thresholds

- Multi-hand tracking support

- Finger state analysis using landmark positions

- Distance measurement between points for gesture classification

The project is implemented with enhancements over basic MediaPipe usage by optimizing the detection parameters and integrating a clear API for easy expansion.

III. Results

Testing was conducted using a standard webcam with the system running at an average frame rate of over 25 FPS, ensuring smooth real-time interaction. The detector accurately identified hand landmarks and finger states under varied lighting and backgrounds. The enhanced drawing utilities provide clear visual feedback.

Compared to existing baseline implementations, the project improved detection speed and provided a modular architecture for gesture extensions.

IV. Conclusion

This paper demonstrates a practical approach to real-time hand gesture recognition using MediaPipe and OpenCV. The implemented system achieves reliable performance and offers a foundation for developing advanced virtual interaction applications. Future work may include integrating machine

learning models to classify a wider range of gestures and enhancing robustness in complex environments.

References

[1] MediaPipe Hands, Google, https://google.github.io/mediapipe/solutions/hands.html

[2] Bradski, G., Kaehler, A., Learning OpenCV, O'Reilly Media, 2008.

[3] Molchanov, P., et al., Hand Gesture Recognition with 3D Convolutional Neural Networks, CVPR,2015.