International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 06 | June – 2025 DOI: 10.55041/ISJEM04137
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

# TOUCHLESS INTERACTION SYSTEM – VIRTUAL MOUSE USING AI

Thasna Nasreen J[1]
[1] MCA (Department of Computer Applications,
Dr. M.G.R. Educational and Research Institute University, Chennai, India)
E-mail: thasnanasreens202217@gmail.com

**Abstract**

This paper describes a Touchless Interaction system that aims to enable people to interact with a wide range of daily-life services in a touchless manner. Gesture recognition is used by the assistant through MediaPipe, voice commands are processed by Google Speech-to-Text allows for hands-free experience. The modular architecture that the system comprises virtual pointing device, speech-to-text processing, voice command processing, and an user interface. The Touchless Interaction System increases accessibility, independence, and hygiene in environments such as education, healthcare, smart environments.

Keywords: Virtual assistant, gesture recognition, voice command, assistive technology, accessibility.

## 1. INTRODUCTION

The touchless interfaces, based on hand gestures and voice recognition, have the advantage of better ergonomics, improved hygiene, and greater accessibility from society aspect **[2]**, **[3]**, **[1]**. Thanks to computer vision libraries such as OpenCV and frameworks for hand detection and tracking such as MediaPipe, gesture recognition has been advancing in recent years making it possible to track hand movements in real time and with high accuracy even from a normal webcam **[5]**, **[4]**. These technologies detect hand landmarks and motion paths to replicate typical input actions (such as cursor movement, clicking, and selecting).

## 2. METHODOLOGY

The following are the core algorithms and techniques used in the development of the system:

### 2.1 Hand Landmark Detection (MediaPipe, OpenCV Hands Algorithm)

**Technique:**

MediaPipe's hand tracking solution uses a two-stage pipeline:

**Palm Detection:** OpenCV captures the palm region from the input frame.

**Hand Landmark Model:** MediaPipe detects and tracks 21 hand landmarks (key points per hand) within the detected palm region.

**Usage in Project:**

Gesture input for cursor control, click events, and virtual keyboard typing.

Specific landmark distances and angles are used to identify gestures like pinch.

 

### 2.2 Natural Language Processing (NLP) for Command Interpretation

**Technique:**

Tokenization and pattern matching using Python's re and nltk libraries.

Keyword and intent extraction from transcribed speech.

**Usage in Project:**

Parse command sentences into executable tasks.

Identify action verbs (e.g., “Launch gesture recognition”).

## 3. MODULE DESCRIPTION

### a) Virtual Mouse Module

The Virtual Mouse Module captures real-time video through a webcam, with MediaPipe detecting **21 hand landmarks to interpret gestures**. These gestures are mapped to mouse functions.

### b) Voice Command Module

The Voice Command Module processes voice input using Google’s Speech Recognition API to interpret spoken command for opening the virtual mouse. The captured voice command is then parsed through **Natural Language Processing (NLP)** to understand the user’s intent.

## CONCLUSION

This project introduces a Touchless Interaction System, by integrating gesture and speech recognition within a modular framework, the system aims to improve digital accessibility and encourage greater independence. Future developments will focus on incorporating predictive intent modeling.

## REFERENCES

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[2] F. Zhang, J. Yang, and D. Zhang, “Hand Gesture Recognition Using Deep Convolutional Neural Networks,” *International Journal of Computer Applications*, vol. 170, no. 2, pp. 29–33, 2017.

[3] D. Grimes and R. James, “A Real-Time Hand Gesture Recognition and Human-Computer Interaction System,” *Procedia Computer Science*, vol. 131, pp. 107–114, 2018.

[4] T. Zhang, X. Yao, and M. Wang, “Vision-Based Hand Gesture Recognition Using Deep Learning for Human–Robot Interaction,” *Sensors*, vol. 21, no. 1, 2021.

[5] MediaPipe by Google, 2025. Online: https://mediapipe.dev