AI-Driven Multimodal Holographic Human Computer Interaction System

Project Report

(Version 1.0)

Course: B.Tech CSE AI ML DL

Degree: Bachelor of Technology (CSE)

PROJECT GUIDE:

MS. VARTIKA GUPTA

SUBMITTED BY:

Y. SIDDHARTH

(TCA2359240)

We hereby declare that this Project Report titled _____ submitted by us and approved by our project guide, Faculty of Engineering & Computing Sciences. Teerthanker Mahaveer University, Moradabad, is a bonafide work undertaken by us and it is not submitted to any other University or Institution for the award of any degree diploma / certificate or published any time before.

Field	Name	Signature
Student Name:	Y SIDDHARTH	
Project Guide :	MS. VARTIKA GUPTA	
Project ID :		

Table of Contents

Appendix

A: Data Flow Diagram (DFD)

B: Entity Relationship Diagram (ERD)

C: Use Case Diagram (UCD)

D: Data Dictionary (DD)

E: Screen Shots

1 Project Title

AI-Driven Multimodal Holographic Human–Computer Interaction System

2 Problem Statement

Traditional human–computer interaction relies heavily on physical input devices such as keyboards, mice, and touchscreens. These interfaces limit natural interaction and are not suitable for environments requiring hands-free or sterile interaction (e.g., healthcare, laboratories, defense). Existing systems lack the ability to provide multimodal interaction combining gesture, voice, and 3D holographic-like visualization.

This project addresses the gap by developing an AI-driven multimodal holographic interaction system that:

- Enables gesture-based and voice-based interaction
- Simulates floating holographic-style 3D UI
- Provides a contactless, intuitive, immersive user experience
- Enhances accessibility and reduces dependency on physical devices
- Uses AI models for real-time recognition, tracking, and system response

Therefore, the problem identified is the absence of an integrated, affordable, and AI-enhanced holographic interaction system capable of supporting next-generation human–computer interaction.

3 Project Description

The system allows users to interact with a holographic-like floating interface using hand gestures, voice commands, and AI-driven tracking. It integrates gesture recognition, voice processing, and holographic UI simulation to create a futuristic multimodal interface. The goal is to provide an intuitive, contactless, and immersive interaction model.

3.1 Scope of the Work

In-Scope:

- Gesture detection
- Voice recognition
- Holographic UI simulation
- Hand tracking
- Floating menu interactions

Out-of-Scope:

• Real hologram hardware
• AR/VR headset integration

3.2 Project Modules

Module 1: Gesture Detection System – Detects gestures using AI and MediaPipe.
Module 2: Voice Command System – Processes user voice inputs using NLP.
Module 3: Holographic UI Simulation – Renders floating 3D-like UI.
Module 4: Interaction Engine – Maps user inputs to UI actions.
Module 5: System Dashboard – Controls system settings and monitoring.

3.3 Context Diagram (High Level)

Inputs:

- Gesture Input
- Voice Commands

System Processing:

- AI Gesture Recognition
- Speech Processing
- Holographic UI Rendering
- Interaction Mapping

Outputs:

- Holographic-style UI Response
- System Action Execution

External Input	System Process	Output
Gesture Input	Gesture Detection Module → Interaction Engine	UI Action Triggered
Voice Command	Voice Processing Module → Interaction Engine	System Response
User Action	Holographic UI Module	Updated UI Display

4 Implementation Methodology

The project follows a structured SDLC approach:
1. Requirement Analysis
2. System Design (DFD, ERD, Architecture)
3. Development of gesture + voice modules
4. Integration with holographic UI
5. Testing and validation

Testing involves:
- Gesture accuracy testing
- Voice recognition accuracy
- UI usability analysis

5 Technologies to be used

Software Platform:
- Front-end: Python UI / WebGL-based simulation
- Back-end: Python AI Models

Hardware Platform:
- 8GB RAM, 256GB Storage
- Webcam
- Windows OS

Tools Used:
- Python
- OpenCV
- MediaPipe
- Google Speech API
- VS Code
- GitHub

6 Advantages of this Project

- Contactless interaction
- Enhanced accessibility
- Futuristic holographic-like interface
- Useful in healthcare, defense, smart homes, automation
- Reduces dependency on physical devices

7 Assumptions, if any

1. The user has a functioning webcam for gesture detection.
2. The system will run on a Windows-based computer with minimum **8GB RAM**.
3. The lighting conditions in the environment will be sufficient for accurate hand tracking.

4. The user will speak voice commands clearly for proper recognition.
5. Stable Python environment and required libraries (OpenCV, MediaPipe, Speech API) are assumed to be installed.
6. Internet connectivity may be required for voice recognition APIs.
7. The system will be operated by a single user at a time.

• 8. Gesture and voice commands used by the user follow predefined formats.

8 Future Scope and further enhancement of the Project

• Integration with AR/VR hardware
• Real hologram projection devices
• AI-powered virtual assistant avatar
• IoT-enabled smart environment integration

9 Project Repository Location

Lab ID: CCSIT-Lab-03
Server Folder: Holo-UI-01
Contains:
• Source code
• Requirements
• Test cases
• Project report
• Diagrams

S#	Project Artifacts (softcopy)	Location (Lab-ID, Server ID, Folder Name)	Verified by Project Guide	Verified by Lab In-Charge
1.	Project Synopsis Report (Final Version)	**CCSIT-Lab-03 / Server-02 / Holo-UI-01**		
2.	Project Progress updates	**CCSIT-Lab-03 / Server-02 / Holo-UI-01**		
3.	Project Requirement Specifications	**CCSIT-Lab-03 / Server-02 / Holo-UI-01**		
4.	Project Report (Final	**CCSIT-Lab-03 /**		

10 Definitions, Acronyms, and Abbreviations

AI – Artificial Intelligence
NLP – Natural Language Processing
HCI – Human–Computer Interaction
UI – User Interface

Abbreviation	Description
AI	Artificial Intelligence
HCI	Human–Computer Interaction
UI	User Interface
NLP	Natural Language Processing
CV	Computer Vision

Abbreviation	Description
DFD	Data Flow Diagram
ERD	Entity Relationship Diagram
UCD	Use Case Diagram
API	Application Programming Interface

11 Conclusion

The project successfully demonstrates a multimodal holographic human–computer interaction system that integrates gesture recognition, voice command processing, and holographic-style UI simulation. The system provides a contactless, intuitive, and futuristic mode of interaction, reducing dependency on physical devices and enhancing accessibility for users.

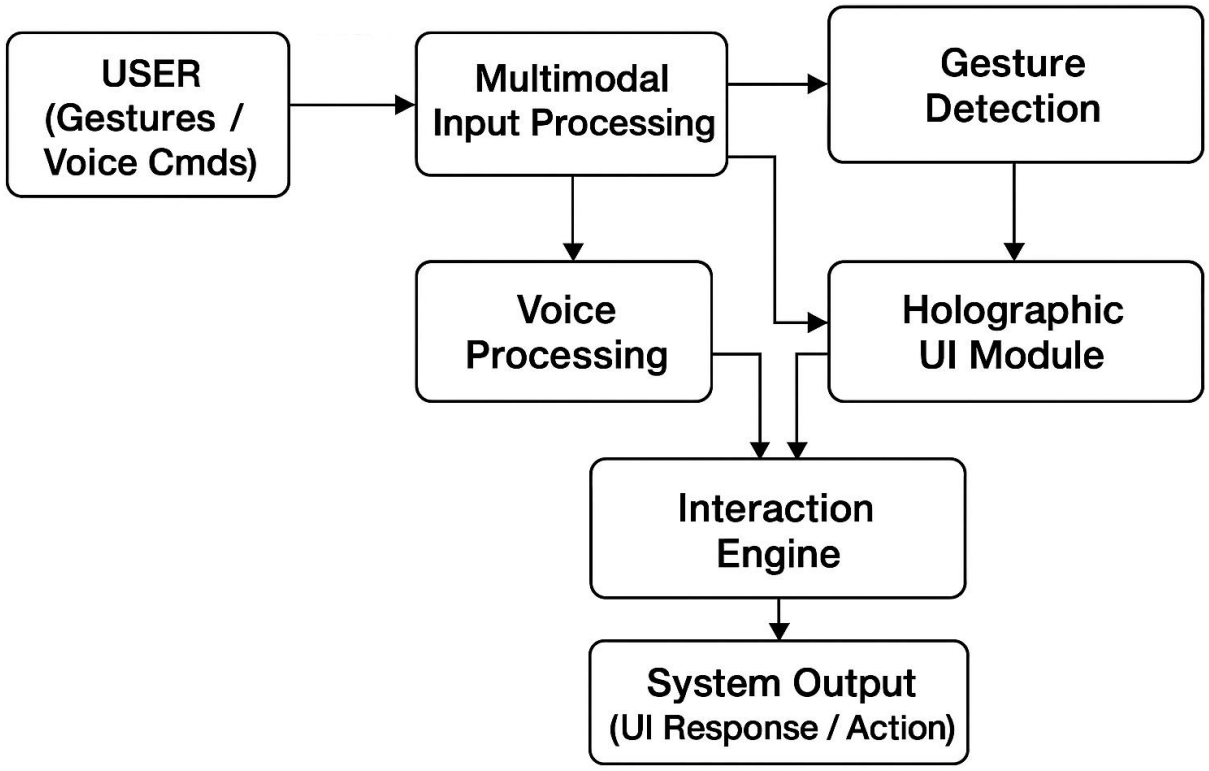
The modular architecture and scalable design allow for easy extension into advanced applications such as AR/VR integration, IoT-based automation, and real holographic projection systems. This work lays a strong foundation for developing next-generation interaction technologies and contributes to the growing field of intelligent user interfaces.

12 References

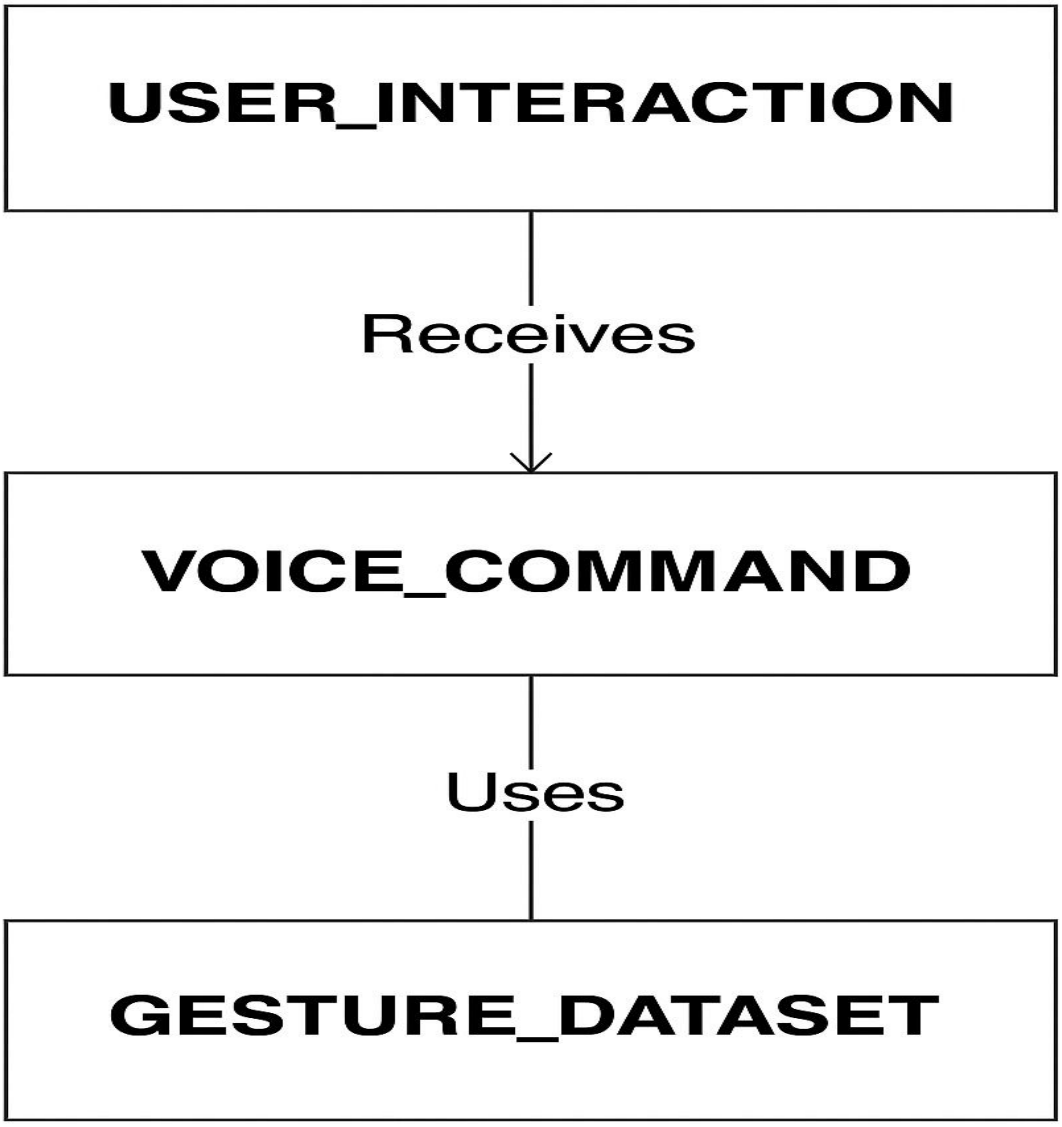
- OpenCV Documentation
- MediaPipe Documentation
- Google Speech API Documentation
- Research Papers on Holographic Interfaces & Gesture Recognition

S#	Reference Details	Owner	Version	Date
1.	Project Synopsis	Holo-UI-01	1.0	
2.	Project Requirements	Holo-UI-01	1.0	
3.	AI-Driven Multimodal Holographic HCI Research Papers	Public Research	—	
4.	Open CV Documentation	OpenCV.org	Latest	
5.	MediaPipe Documentation	Google	Latest	
6.	Google Speech Recognition API Documentation	Google	Latest	

Annexure A
Data Flow Diagram (DFD)
(Mandatory)



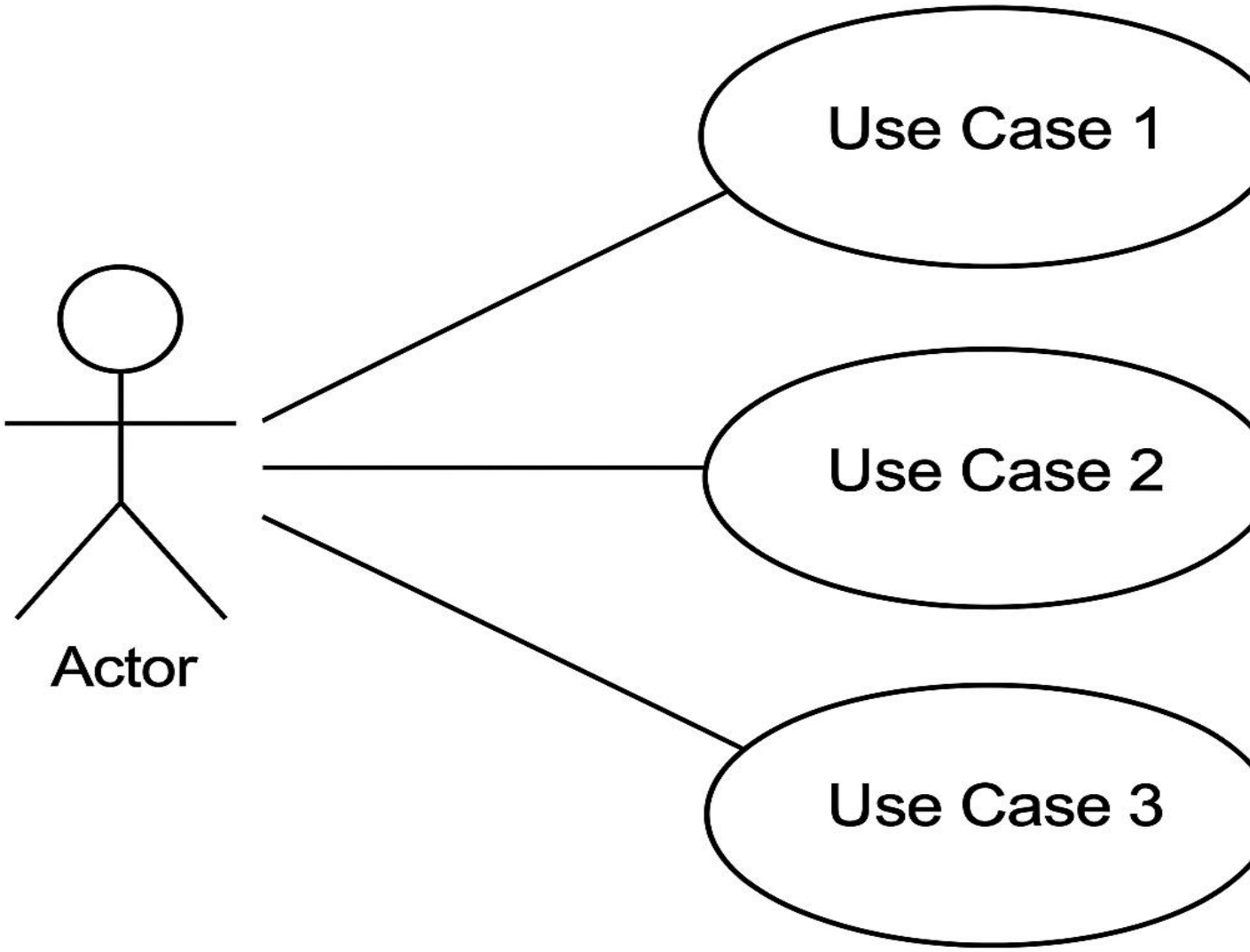
Annexure B
Entity-Relationship Diagram (ERD)
(Mandatory)



Annexure C

Use-Case Diagram (UCD)

(Optional)



Annexure D
Data Dictionary (DD)
(Mandatory)

Example:
User Table (USR)

Fields	Data type	Description
INT-ID	Number	Unique interaction ID
INT-Type	Text	Gesture or Voice
INT-Command	Text	Detected user command
INT-Timestamp	Text	Time of interaction
INT-Response	Text	Output generated by the system

Supplier Table (SUPP)

Fields	Data type	Description
GST-ID	Number	Unique gesture ID
GST-Name	Text	Gesture name (e.g., Swipe Left)
GST-Landmarks	Text	MediaPipe landmark coordinates
GST-Category	Text	Navigation / Action / Selection

Fields	Data type	Description
CMD-ID	Number	Unique voice command ID
CMD-Text	Text	Raw voice input detected
CMD-Intent	Text	Interpreted system intent
CMD-Confidence	Number	Confidence score from recognition model

Annexure E
Screen Shots

Home Page: