universidade de aveiro  theoria poiesis praxis

Multimodal Interaction Using Gestures

IKEA Website Controlling

Roberto Rolão de Castro - 107133
Tiago Caridade Gomes - 108307

Course: Multimodal Interaction 2024/25

Contents

1 Introduction **3**
1.1 Objectives 3
1.2 Application Selection 3

2 System Architecture **3**

3 Gestures Overview **4**

4 Conclusion **5**

List of Figures

2.1 Architecture Diagram 4

1. Introduction

This report outlines the development of an **interactive application controlled exclusively through gestures**, implemented using **Microsoft Kinect**. The primary objective is to explore gesture detection technologies to control diverse functionalities on the IKEA website, enabling intuitive and accessible control.Gestures allow users to interact seamlessly with the website, offering an innovative alternative to traditional input methods.

1.1 Objectives

- Utilize Microsoft Kinect to recognize gestures as the sole input modality.
- Ensure the system allows control of various application functionalities.
- Implement appropriate feedback via speech synthesis.

1.2 Application Selection

The chosen application is based on a previous practical project that employed voice control for interaction with an online shopping system (e.g., IKEA). The scope was adapted for gesture control.

2. System Architecture

The system's architecture integrates several components to enable seamless gesture-based interaction:

- **Microsoft Kinect:** The primary hardware for capturing user gestures. Kinect's depth-sensing cameras and motion tracking capabilities allow accurate recognition of predefined gestures and provide the foundation for gesture-based interaction.
- **Multimodal Interaction Manager (MMI):**Responsible for receiving gesture commands from the Kinect and distributing them to the appropriate system components. It acts as the intermediary between gesture recognition and the application's functionalities.
- **Main.py:** This Python component acts as the central controller, managing the flow of data and coordinating the interaction between different components. It receives commands from the MMI via WebSockets and interacts with the IKEA website via Selenium WebDriver to execute actions such as scrolling and clicking.
- **Selenium WebDriver:**Used for automating interaction with the IKEA website. Selenium executes actions like scrolling, clicking buttons, and navigating product pages based on gesture inputs.
- **WebAssistantApp:**A web application that serves as the interface layer, enabling communication between the user and the system. It provides visual feedback and supports interaction workflows.

- **Speech Interface:**Although this project focuses on gestures, the speech interface complements the interaction by providing feedback through synthesized voice outputs, ensuring users are informed about system actions and statuses.

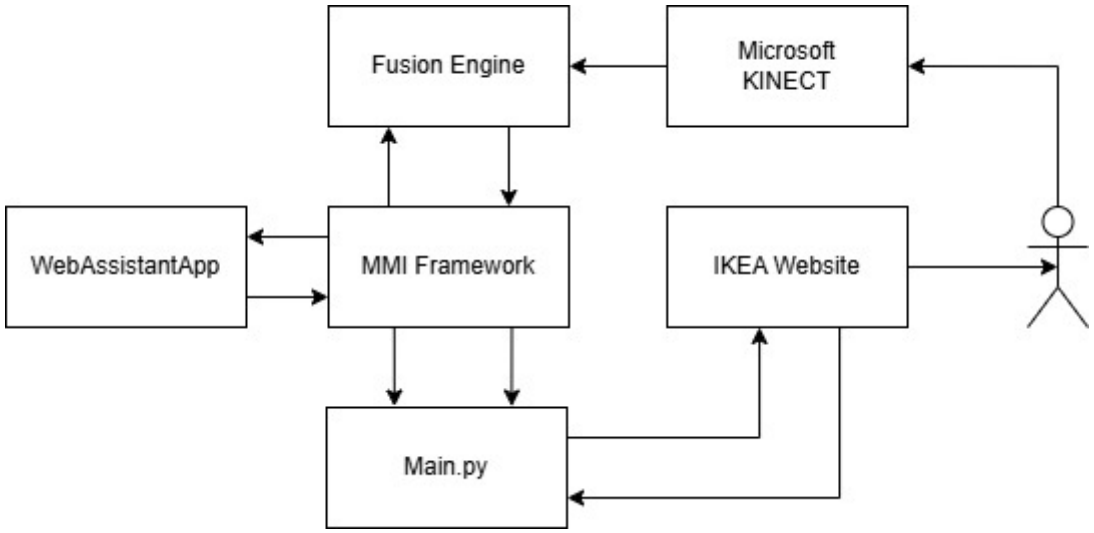

Figure 2.1: Architecture Diagram

This architecture ensures modularity and scalability, allowing components to operate independently while working together to deliver a cohesive user experience. Each module is optimized for its specific function, from gesture recognition to website interaction.

3. Gestures Overview

The system implements the following gestures to facilitate interaction with the IKEA website:

- **Open Site:** Perform a gesture with closed hands extended to the left to open the IKEA Portugal website.
- **Add to Cart/Favorites:**
  - Add a specific product to the cart by raising the right arm with the hand open.
  - Add a specific product to the favorites list with both closed hands extended forward.
- **Scroll the Page:** Extend the arm to the side-bottom with half hand closed to scroll the page (right-up and left-down).
- **Search:** Perform a gesture with open right hand extended forward to search for a product on the website.
- **Go Back:** Raise the left hand in direction to the shoulder to go to the previous page.
- **Main Page:** Raise the right hand in direction to the shoulder to return to the homepage.

This gesture-based interaction enables users to navigate and control the IKEA website efficiently, enhancing accessibility and providing a unique user experience.

4. Conclusion

The development of this gesture-based interaction system highlights the potential of multimodal interfaces in improving user accessibility and experience. By utilizing Microsoft Kinect and integrating it with the IKEA website, the project successfully demonstrates the effectiveness of gestures as a primary input modality. The system's architecture, combining gesture recognition, automation through Selenium WebDriver, and interactive feedback, ensures smooth operation and user satisfaction.