HCI - MultiModal Systems - Theory 2

Artificial Intelligence

Jacopo Parretti

I Semester 2025-2026

Indice

Parte I

Foundations of Multimodal Interaction

1 Introduction to Intelligent Multimodal Interfaces

1.1 Course Objectives

This course explores the fundamental theories and concepts of **Human-Computer Interaction (HCI)**, an interdisciplinary field that synthesizes knowledge from cognitive psychology, computer science, and design. The primary objectives are:

- Understanding the theoretical foundations of human-computer interaction
- Developing practical skills in designing and implementing multimodal interfaces
- Exploring the integration of artificial intelligence techniques in interactive systems
- Analyzing nonverbal communication and its role in human-computer interaction

1.2 Focus Areas

The course places special emphasis on three interconnected dimensions:

1.2.1 Technological Solutions

Students will develop computer interfaces with a focus on both methodological and implementation aspects. The course emphasizes hands-on experience in building functional interactive systems, bridging the gap between theory and practice.

1.2.2 Multimodal Interaction

Special attention is devoted to **multimodal solutions** that integrate multiple input and output modalities:

- **Touch**: Tactile and haptic interaction
- **Vision**: Camera-based interaction and visual recognition
- **Natural Language**: Speech and text-based communication
- **Audio**: Sound-based interaction and auditory feedback

1.2.3 Intelligent Systems

The course explores how **artificial intelligence techniques** can enhance interaction by:

- Inferring user intentions from multimodal input
- Predicting expected interactions based on context and user behavior
- Adapting interface behavior to individual users
- Recognizing and responding to affective and social cues

1.3 Course Program

1.3.1 Theoretical Component

The theoretical component covers the following topics:

1. **Introduction**: Course motivation, professional perspectives, open research issues, program overview, and examination methodology
2. **Foundations of HCI**: Human factors in interface design, interaction design principles, usability evaluation, gaming, and gamification
3. **Visual Interaction**: Camera calibration techniques, structure from motion, 3D reconstruction
4. **Nonverbal Behavior in Communication**:
   - Types of nonverbal behavior: facial expressions, gestures, posture, eye gaze
   - Data collection methods and protocols
   - Tools and software for nonverbal behavior analysis
   - Annotation tools (e.g., ELAN)
5. **Automated Analysis of Body Language**: Movement tracking, gesture recognition, facial expression analysis, and speech processing. Techniques for data capture, feature extraction, and automatic analysis
6. **Social Artificial Intelligence**: Applications in social psychology, organizational psychology, and social robotics
7. **Affective Computing**: Theories of emotion, emotion recognition systems, and applications in HCI
8. **Multimodal Fusion**: Integration of multimodal nonverbal cues using fusion techniques (late fusion, early fusion)

1.3.2 Laboratory Component

The laboratory sessions provide hands-on experience with state-of-the-art tools and techniques:

1. **Deep Image Matching**: Python implementation of feature detection and matching algorithms
2. **3D Model Reconstruction**: Structure from motion using Zephyr software
3. **Camera Pose Estimation**: C# implementation of Fiore's method for camera localization
4. **3D Graphics**: Modeling and rendering in Unity game engine
5. **Model-Based Augmented Reality**: Implementation of the complete AR pipeline integrating Python code and Unity
6. **Advanced Topics**: Deep learning approaches to camera pose estimation and model recognition