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

2 The Evolution of Human-Computer Interaction

2.1 Why Study HCI?

The fundamental motivation for studying Human-Computer Interaction lies in understanding the **interaction space** between humans and machines. Effective interaction design must strike a balance between these two poles—neither overly machine-centric nor purely human-centric, but rather positioned at an optimal middle ground that accommodates both human capabilities and computational constraints.

The central challenge of HCI is to design interfaces that allow users to accomplish their goals efficiently and naturally, while leveraging the computational power and capabilities of modern computing systems.

2.2 Historical Evolution of Interaction Paradigms

The field of HCI has undergone several paradigm shifts, each bringing the interaction closer to natural human behavior.

2.2.1 Command-Line Interfaces (Pre-1980s)

Early human-computer interaction relied on **command-line interfaces (CLI)**, which required users to communicate with machines through text-based commands. This paradigm represented an interaction model heavily biased toward the machine:

- **Expert-oriented**: CLIs assume users possess detailed knowledge of the system's internal organization and command syntax
- **Machine-centric**: The interface reflects the machine's architecture rather than human cognitive models
- **High learning curve**: Users must memorize commands, parameters, and system-specific conventions
- **Limited accessibility**: Generic users cannot interact effectively without substantial training

While command-line interfaces persist today (particularly among expert users who value direct system control and automation capabilities), they exemplify an interaction paradigm positioned closer to the machine than to the human.

2.2.2 Graphical User Interfaces (1980s–2000s)

The introduction of **Graphical User Interfaces (GUIs)** in the early 1980s marked a fundamental shift in interaction design:

- **Direct manipulation**: Users interact with visual representations of objects through pointing devices (mouse) and keyboards
- **Visual metaphors**: Desktop metaphors, windows, icons, and menus provide intuitive representations of system functionality
- **Reduced cognitive load**: Users no longer need to memorize complex command syntax
- **Increased accessibility**: Non-expert users can accomplish tasks through exploration and recognition rather than recall

This paradigm shift moved interaction significantly closer to human cognitive models, making computing accessible to a broader population.

2.2.3 Pervasive and Ubiquitous Interaction (2000s–Present)

Over the past 15–20 years, a new paradigm has emerged: **pervasive interaction** (also known as ubiquitous computing). This paradigm is characterized by:

- **Distributed computing**: Intelligence is distributed across multiple interconnected devices (smartphones, wearables, IoT sensors, smart home devices)
- **Invisible interfaces**: The traditional notion of "the machine" becomes diffuse—computation is embedded in the environment
- **Context awareness**: Systems adapt to user context, location, and activity
- **Natural interaction**: Voice, gesture, and multimodal input replace traditional keyboard and mouse
- **Smart environments**: Individual devices collaborate to create intelligent, responsive spaces

In contrast to earlier paradigms where "the machine" was a discrete entity contained in a single case, pervasive computing dissolves the boundaries between user, environment, and computational infrastructure. This represents the closest approximation yet to natural human interaction patterns.

2.3 Technological Foundations

This course emphasizes the **technological aspects** that enable modern HCI systems. Implementing sophisticated, intelligent, and multimodal interactions requires substantial infrastructure:

- **Network infrastructure**: High-bandwidth internet connectivity, satellite networks, and edge computing for distributed systems
- **Interaction devices**: Advanced input/output hardware including VR/AR headsets, depth cameras, haptic devices, and multimodal sensors
- **Computational power**: GPU acceleration, cloud computing resources, and specialized AI hardware for real-time processing
- **Software frameworks**: Machine learning libraries, computer vision toolkits, and multimodal fusion platforms

Understanding these technological foundations is essential for designing and implementing the next generation of human-computer interfaces.

3 Human Factors