48727 Fall 2024 Final Project Proposal

Author: Chia Hui Yen

Prof. Vernelle A. A. Noel, Ph.D. ([vnoel@andrew.cmu.edu](mailto:vnoel@andrew.cmu.edu)) ; TA: Stella Shen ([chenyis@andrew.cmu.edu](mailto:chenyis@andrew.cmu.edu))

**SenseBridge: Enhancing Hybrid Communication Through Motion-Aware Interaction**

# ****Preface****

This project explores how non-verbal cues, captured through body motion, can bridge the gap between virtual and physical communication. Building on the ubiquity of computational systems (ubiquitous computing), it integrates human motion as invisible data, revealing emotional nuances to enhance hybrid living. Inspired by Pierre Lévy’s assertion that virtualization “creates new realities by introducing irreversibility, indeterminacy, and creativity” (Becoming Virtual, p. 29), this project reimagines motion as a universal language that connects human agents across physical and virtual domains.

This work builds on foundational ideas introduced in previous assignment in the course, Assignment A6: Hybrid Bodies, which explored the integration of virtual information into daily life. By critically examining hybrid living as a future norm[[1]](#footnote-1) , the project focuses on reimagining how users engage with and approach emotional communication in hybrid environments—emphasizing medium, timing, and context.

## ****Concept and Central Questions****

1. **What constitutes an effective system for hybrid interaction?**
2. **How can non-verbal cues, such as body motion, enhance emotional communication?**
3. **How can motion data serve as a bridge between physical and virtual spaces, preserving authenticity?**

By addressing these questions, the project highlights the balance between human-centric design and technical complexity, advancing hybrid communication to new levels of immersion and connectivity.

**Theoretical Framework that related to my proposal:**

1. **Virtualization as a Creative Force (Lévy):** Virtualization transcends the “here” and “now,” creating continuity without spatial unity (p. 31). This principle underpins the design of asynchronous 3D messaging and dynamic video interactions that adapt to users’ emotional states.
2. **The Medium is the Message (McLuhan):** Body motion becomes the medium, shaping the message itself and transforming virtual interactions into a more humanized experience. McLuhan’s insights drive the project’s prioritization of the medium over traditional text-based or video communication.
3. **Hybrid Spaces and Tangible Interaction:** The duality of physical and virtual gestures mirrors the hybrid model of modern living. This allows users to engage in richer, more intuitive interactions in both synchronous (video calls) and asynchronous (3D messaging) contexts.

# ****Prototype: SenseBridge****

## ****System Workflow and Features****

The system workflow involves three core components: **Input** (data capture), **Processing** (semantic understanding), and **Output** (visualization and interaction).

#### **1. Input: Data Capture**

* Use a normal webcam (on-call) or Kinect sensor (off-call) to track motion cues such as posture, walking speed, and hand gestures.
* Encode subtle or intentional cues into emotional states using classification algorithms.
  + Supporting Research: Behavior reflects one’s working status (Sarver et al., 2015).

#### **2. Processing: Semantic Understanding**

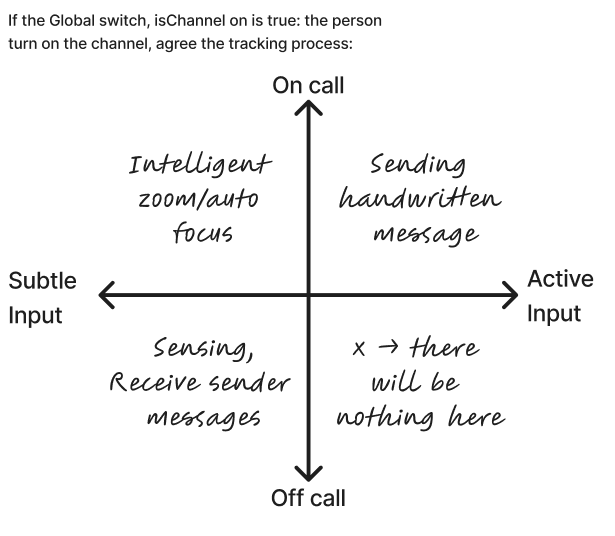
* **Tools**: TensorFlow or PyTorch to analyze and generalize motion data.
* **Workflow**:
  + Collect dynamic motion patterns and map them to emotional states.
  + Example: Walking speed and frequency of movement indicate mood.
  + Train using labeled datasets; test and refine for accuracy.
  + Supporting Concept: “Interaction between humans and systems implies a dialectic between the virtual and the actual” (Lévy, p. 27).

#### **3. Output: Visualization and Interaction**

Translating the processed data into visual, auditory, or tactile feedback, including animations, 3D messages, or dynamic adjustments in video calls.

## Mapping Features into 4 Groups

* The system maps all features into four categories based on two scales:
  + **OnCall/OffCall**: Whether the feature operates during active communication (on-call) or passively in the background (off-call).
  + **Subtle Input/Intentional Input**: Whether the user captures the motion data unintentionally (subtle) or deliberately (intentional).



## ****Exact Feature Description****

Demo can be found in the attached powerpoint (with video).

**3D Messenger, Virtual Sticky Notes (OnCall, Intentional Input)**

* **Description**: The 3D messenger allows users to create and share virtual sticky notes by drawing patterns in the air. Motion data is captured using Kinect sensors to track hand gestures. The system records the user’s hand trail and converts it into a 3D model.
  + **How It Works**:
    - The sender creates a 3D message by performing deliberate gestures in space.
    - The sender specifies a location and pose (e.g., standing in front of a fridge) for message delivery.
    - The message appears when the recipient matches the location and pose, triggering an animation or effect.
  + **Use Case**: This feature mimics the experience of leaving physical sticky notes but enriches it with digital interactions. It enhances asynchronous communication by tying messages to specific spatial contexts.

**Sensing (OffCall, Subtle Input)**

* **Description**: The sensing feature passively monitors body motion to infer emotional states without active user interaction. Subtle cues such as walking speed, movement frequency, and sitting patterns are analyzed to reflect the user’s status.
  + **How It Works**:
    - Motion data is continuously captured via Kinect sensors.
    - Machine learning models classify the collected data to identify emotional or physical states (e.g., stress, calmness).
    - This data can be visualized or used as background insights to inform interactions.
  + **Use Case**: For long-distance communication, sensing can provide family members or partners with subtle clues about each other’s moods, fostering emotional understanding without explicit messaging.

**Video Call Enhancer (OnCall, Subtle Input)**

* **Description**: The video call enhancer improves real-time communication by dynamically adjusting visuals and audio based on user behavior.
  + **Dynamic Background Music Adjustment**:
    - The system detects when a user is speaking and automatically reduces background music volume to ensure clear communication.
    - When the user stops speaking, the music volume gradually increases to create a balanced experience.
  + **Hand Gesture Zoom**:
    - Deliberate hand gestures (e.g., waving) trigger the system to zoom in on the user, emphasizing active participation and engagement.
  + **Distance-Based Zoom**:
    - The system calculates the user’s distance from the camera using motion landmarks. If the user moves further away, the system zooms in to maintain visibility. If the user is close to the camera, the view remains stable to avoid overcorrection.
  + **Use Case**: This feature simulates the attentiveness of in-person conversations, where physical proximity and gestures influence engagement. It is particularly useful for enhancing remote meetings, virtual social gatherings, or long-distance interactions.

## ****Impact and Evaluation****

#### **Evaluation Metrics:**

1. **Accuracy:** Motion detection accuracy (e.g., sensitivity, specificity).
2. **Usability:** User testing to evaluate intuitiveness and engagement.
3. **Immersion:** Surveys assessing participants’ sense of presence and connection.

#### **Critical Potential:**

The project challenges conventional communication methods by leveraging XR to facilitate meaningful, hybrid interactions. By integrating motion-based data, the system enhances emotional connection while avoiding excessive, unnecessary information. Lévy’s concept of escaping spatial constraints (p. 31) is key to understanding this transformative potential.

# ****Conclusion****

This project showcases the power of XR in humanizing virtual interactions through motion-based communication. By leveraging non-verbal cues, it bridges the emotional and practical gaps of hybrid living, redefining the boundaries of virtual and physical worlds. It provides a technical foundation and creative inspiration for future exploration into hybrid communication systems.

# ****References****

1. Sarver, D.E., Rapport, M.D., Kofler, M.J., et al. (2015). Journal of Abnormal Child Psychology, 43, 1219–1232.
2. Lévy, Pierre. Becoming Virtual: Reality in the Digital Age. Plenum Trade, 1998.
3. McLuhan, Marshall. Understanding Media: The Extensions of Man. McGraw-Hill, 1964.
4. https://www.e-flux.com/architecture/intelligence/310403/too-much-information/.

1. Becoming Virtual, Pierre Lévy, Proof the “actualness” of virtual world:

   “If virtualization were nothing more than the transition from a realitxcy to a col¬ lection of possibles, it would be derealizing. But it im¬ plies as great a sense of irreversibility in its effects, indeterminacy in its processes, and creativity in its striv¬ing, as actualization.”

   “Virtualization is one of the princi¬pal vectors in the creation of reality.” [↑](#footnote-ref-1)