POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

Bridging Realities: Advancing Human-Avatar Interaction through Sensory Translation Systems in the Physical Metaverse

Laurea Magistrale in Computer Science and Engineering - Ingegneria Informatica

Author: Maurizio Vetere

Advisor: Prof. Andrea Bonarini

Co-advisor: Federico Espositi

Academic year: 2022-23

1. Introduction

This thesis addresses the challenge of embodying non-humanoid Physical Avatars to interact with a human Visitor. It presents a Sensory Translation System connecting a human Controller with the Physical Avatar, to explore nonverbal communication in a structured goal-oriented activity with the Visitor. The research incorporates VR headsets, robots with sensory inputs, and software for environmental information and human pose translation, which elaborates camera information to mediate communication between Avatar and Human. Key experiments include the "First Contact" experience and the "Mazes" scenarios, demonstrating the technology's potential in human-robot interaction and contributing to the Physical Metaverse concept.

2. Background

The project evolves from the concepts presented in the pilot paper *Towards a Framework for Embodying AnyBody Through Sensory Translation and Proprioceptive Remapping: a Pilot Study* [2]. It aims to translate theoretical insights into practical software and hardware applications. This includes the integration of an initial Sensory Translation System developed in a previous thesis [1], adapted to work with new robotic platforms [3]. Our project was presented in the events "Playful Machines" at Milano Digital Week and "First Contact" at the X-Cities exhibition. We used them as opportunities to collect experimental data from real users, showcasing the potential of this technology in validating the Physical Metaverse framework.

3. Theoretical Framework

We laid down the theoretical foundations required to introduce the newly added goal-oriented activity to our study in the Physical Metaverse context, focusing on selecting a suitable one for interaction between the Physical Avatar and the Visitor. The goal was to find a non-verbal, two-player goal-oriented activity to test Sensory Translation and Human Translation components [Figure 1]. Various physical and cognitive games were analyzed based on their relevance to the system's needs, such as information transfer from Avatar to Visitor and Visitor to Avatar, the Avatar's proprioception, and the required environment knowledge. The Escape Room activity was chosen for its balance of project requirements and practical con-

straints. We also discussed about the virtualization of elements like the Visitor and the Avatar for system testing and development purposes.

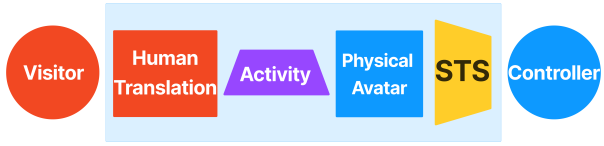


Figure 1: Graphical representation of our system. STS stands for Sensory Translation System.

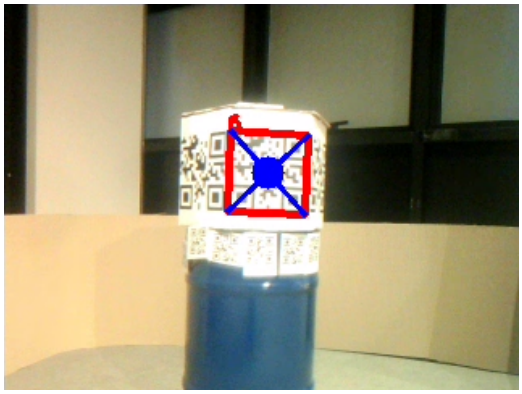
4. Project Implementation

With the implementation of this project, we observe a fascinating journey of integrating advanced technologies to create a seamless interface between the physical and digital realms. The project, grounded in the concept of Sensory Translation within the Physical Metaverse, demonstrates a sophisticated synthesis of electronic devices, software solutions, and interactive environments. This endeavor is not just a technological feat but a vivid illustration of how theoretical concepts can be translated into tangible experiences.

A key technical focus of the project is Sensory Translation, primarily revolving around visual perception. This aspect is ingeniously addressed through the use of a VR headset that provides the Controller with a first-person perspective of the Physical Avatar's environment. The integration of LIDAR technology is a strategic technical choice, enhancing the system's ability to navigate and interact with its surroundings. The implementation of the DepthAI camera is particularly noteworthy. It represents a significant technical upgrade from the initial setup, enabling more accurate and efficient human pose estimation [Figure 2a].

The technical evolution from the project's initial foundation is evident in the systematic enhancement of existing features and the introduction of new components. The refinement of the LIDAR visualization, for instance, exemplifies the project's approach of building on previous work. By manipulating color and light in the virtual environment, for example with the addition of a sun in the sky and a floor texture on the ground, the developers were able to enhance the stability

(a) Pose detection.



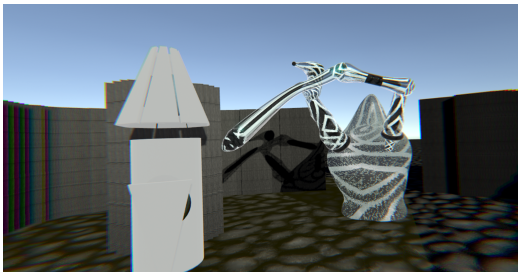
(b) Station detection.

Figure 2: Camera detections with pose and QR recognition.

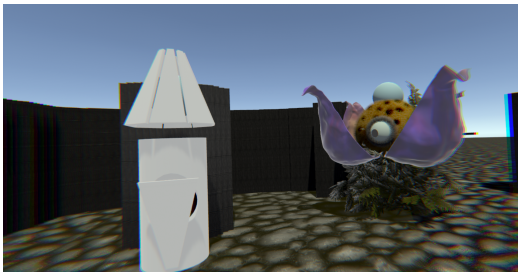
and clarity of the visualization. This improvement not only contributes to the visual quality of the virtual environment but also aids in better spatial orientation and navigation for the Controller.

QR code Station tracking introduces another layer of technical sophistication. By detecting QR codes attached to the physical objectives of our Escape Room activity [Figure 2b], namely the Stations, the system can now identify and allow interaction with relevant objects in the environment. This feature is critical for the Controller to extract useful information for navigation and interaction. The incorporation of haptic and audio feedback upon interaction with the Stations is a noteworthy enhancement to our Sensory Translation, expanding the range of sensory channels we translate.

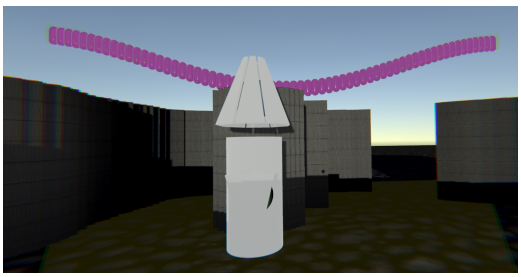
This system focused on three different Human Translations of our Visitor: Odile, Siid and Evangelion. The first two are inspired from robots of the Physical Metaverse, while the third one recalls one of the previous work's Human Translations [Figure 3].



(a) Odile.



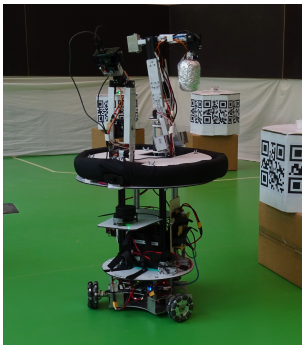
(b) Siid.



(c) Evangelion.

Figure 3: Sensory Translation System visualization of the detections from Figure 2. The three different Human Translations we developed are shown.

The implementation on Odile and the subsequent switch to Blackwings illustrate a flexible and adaptive technical strategy. The modification of these robots to accommodate new hardware, including the pan and tilt mechanism of the camera, highlights the project's technical ingenuity.
The creation of a digital Simulation for our system [Figure 5] marks a significant technical milestone in the project. By providing a digital twin of the real system, we were able to efficiently test and refine the Sensory Translation System with more agility. This virtual environment played a crucial role in isolating hardware-related complications, allowing the team to focus the efforts more efficiently on software challenges.
The public engagements during Digital Week and the X-Cities Exhibition not only served as



(a) Odile.

(b) Blackwings.

Figure 4: Odile and Blackwings with Stations in the background.

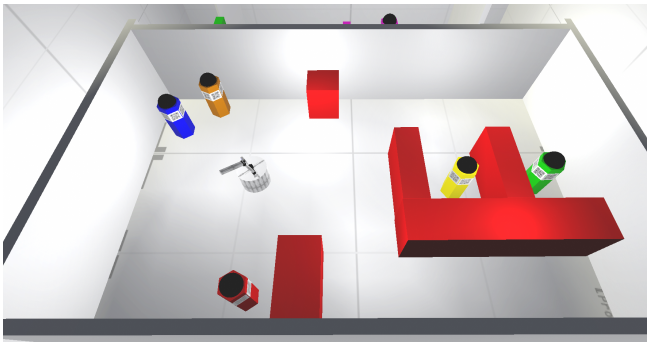


Figure 5: Simulation of our system. The white cylinder with an arm is Virtual Odile, the prisms are Stations.

demonstrations of the system's capabilities but also as critical phases for technical refinement. The feedback received from real users informed further technical adjustments, highlighting the importance of user-centered design in technology development.
In summary, the technical narrative of this project's implementation is one of continuous innovation, adaptability, and user-centric design. From leveraging advanced sensors and cameras to developing a comprehensive virtual simulation, the project exemplifies a deep understanding of the technical complexities involved in bridging the physical and digital worlds. The ability to integrate feedback from real-world applications into the technical development process further underscores the project's commitment to creating a robust, responsive, and immersive Physical Metaverse.

5. The Experiments

In the final phase of experimentation at the X-Cities exhibition, the integration of virtual reality, robotics, and Sensory Translation proved successful, creating a compelling blend of physical and virtual elements. The experiments underscored the system's ability to facilitate navigation and foster interaction in a shared virtual and physical environment, through the dynamics of Human-Avatar interaction.
Key Findings:

- Participants effectively engaged through both the virtual and physical components of the project.
- The system demonstrated promising capabilities in facilitating nonverbal communication through non-humanoid Avatars.
- Insights into the dynamics of Human-Avatar interaction were gained.

6. Conclusions

The project outcomes indicate a successful integration of advanced technologies in the Physical Metaverse context. Participants showed meaningful engagement, and the system facilitated an immersive experience, blending virtual and physical interactions. The study also opened new avenues in understanding nonverbal communication through different forms of avatars, with implications across various domains, including telepresence, entertainment, and gaming.

References

[1] Giuseppe Epifani. Non-verbal communication through a robotic physical avatar: A study using minimal sensor information. Master's thesis, School of Computer Science and Engineering - Ingegneria Informatica, Politecnico di Milano, 2022. Student ID: 10578541.

[2] Federico Espositi and Andrea Bonarini. Towards a framework for embodying anybody through sensory translation and proprioceptive remapping: a pilot study. *To be Published*, 2023.

[3] Erica Panelli. Odile: An expressive robotic agent for emotional exchange and information sharing. Master's thesis, School of Computer Science and Engineering - Ingegneria Informatica, Politecnico di Milano, 2023. Student ID: 953584.