E-avatar : Hybrid education through Pose estimation and Mixed reality

Dounia Bougamza[1], Fatima-Zahra Kaghat[2], Ekkarat Rodthong[3], and Ahmed Azough[4]

[1] ESILV Engineering School, Paris La Défense, France
dounia.bougamza@edu.devinci.fr
[2] Léonard de Vinci Pôle Universitaire, Research Center Paris La Défense, France
fatima-zahra.kaghat@devinci.fr
[3] IIM Digital School, Paris La Défense, France
ekkarat.rodthong@devinci.fr
[4] Léonard de Vinci Pôle Universitaire,Research Center Paris La Défense, France
ahmed.azough@devinci.fr

Abstract. Recent world events has led remote working to become a widespread means of working throughout universities and industries. Its over-the-night popularization did not come without a set of limitations. Studies reveal concerns regarding student presence, engagement and motivation. Tools for videoconferencing, which are frequently used for remote work, do not enable realistic interaction between participants. This results in a barrier between the teacher and the distant pupils in the setting of distance learning, as well as a loss of interest on their part. To remedy these limitations, this study proposes a new mode of interaction between teacher and remote students. This mode combines the techniques of mixed reality, pose estimation, holoportation and avatar generation to provide an audiovisual and gestual interaction between teachers and students. Using this new technology, a teacher putting on a mixed reality headset will be able to see and communicate with the avatars of their pupils in real time, placed side by side with the students physically present. We believe that this new mode will encourage collaboration and engagement between teachers and their remote students.

Keywords: Avatar · hybrid learning · pose detection · hologram · unity3D · immersive experience

1 Introduction

According to recent results, student performance in the French education system has significantly declined[5]. Experts attribute this drastic drop in performance

[5] French secondary school students' results drop post-pandemic, as elsewhere (2023):https://www.lemonde.fr/en/france/article/2023/12/05/french-secondary-school-students-results-drop-post-pandemic-as-elsewhere_6314516_7.html#

to the COVID-19 pandemic, which disrupted the educational system through the widespread implementation of remote learning. This new mode of teaching offers numerous advantages through the use of strategies based on audio, video and text. However, it remains limited. The difficulty in maintaining motivation and the lack of face-to-face social interaction have led to an increase in school dropouts. In France, measures such as the "Europe 2020 Strategy" have been implemented by the national education system, with the aim of combating this phenomenon[6]

Despite the persistence of remote learning, the idea of establishing a new distance education system aimed at promoting student engagement has emerged [1]. Furthermore, this can be a useful tool for both the corporate and educational sectors. For instance, video conferencing meetings facilitated by platforms like Teams could simulate face-to-face interactions while operating remotely. This paper proposes a framework for creating an immersive experience where participants appear to be physically present, even when they are physically distant. This allows individuals to see, hear and interact with the hologram as if they were present in the same room (Figure 1). This novel tool has potential applications accross various areas, such as remote business meetings, medical telepresence, distance education and other situations where real-time virtual communication with a more immersive presence is desired.

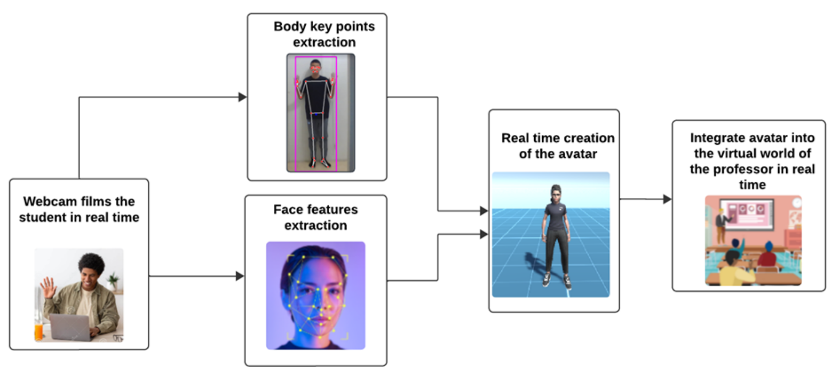

Fig. 1. The development steps of the solution

This article is organized as follows: We cite earlier studies on the application of mixed reality and holoportation in the field of distance communication in Section 2. The elements of our mixed reality-based learning framework are covered in Section 3. A few samples of the work that has been done are shown. In conclusion, Section 4 makes recommendations for future study directions.

---

[6] MNE: https://www.education.gouv.fr/la-lutte-contre-le-decrochage-scolaire

2 Related work

Mixed reality is created by using real images and sounds supported by virtual images and sounds (Billinghurst & Kato, 1999). In this technology, virtual objects are projected into real audio-visual environments which can be observed and interacted with. Numerous studies have explored its potential across different domains, including entertainment, military training, and healthcare [1] [3].
In educational settings, mixed reality holds particular promise. By integrating virtual and real elements, students can engage in more interactive learning experiences. For example, virtual labs can offer practical experimentation in subjects such as chemistry or biology without the risks or limitations associated with physical labs [6].
The goal in our concept is to leverage this by creating a hybrid educational environment where learners equipped with Hololens 2 or Meta Quest 3 headsets can engage in holoported experiences and interact with both real and virtual objects, promoting deeper engagement and understanding.
Holoportation is a technology that combines virtual reality (VR) and 3D capture to allow a person to appear as a hologram in another location in real time [5]. NASA holoported Dr. Josef Schmid, a flight surgeon, from Earth to the International Space Station in October 2021 as a demonstration of this technology[7]. This allowed ESA astronaut Thomas Pesquet to engage in a bidirectional conversation with them, with live images. Because of this technological achievement, human communication over great distances has become possible, opening up new avenues for exploration in space.
In addition, Holoportation, developed by HoloForge, revolutionizes remote communication by surpassing the limitations of traditional audio calls and video conferences. Nevertheless, the user experience may be impacted by its hardware dependencies, particularly Azure Kinect cameras, as well as possible connectivity problems like latency or bandwidth restrictions [2]. These connectivity issues may result in data transmission delays or communication interruptions, thereby reducing the effectiveness of holoportation in unstable or unreliable network conditions. It's essential to consider these limitations for effective implementation and broader adoption[8].
Despite notable advancements in holoportation technology within the domain of remote communication, its adoption within the educational sector is not widespread.

3 The Framework

We propose an innovative solution that fosters an immersive communication experience, akin to face-to-face interaction, by removing the screen barrier and enabling natural exchanges, including non-verbal cues. It offers unlimited scal-

[7] NASA:`https://www.nasa.gov/humans-in-space`
[8] Holoforge:`https://www.holoforge.io/en/`

ability in terms of participants, making it applicable in various industries such as education, collaborative work, and entertainment.

Fig. 2. Teacher's perspective, wearing the Hololens headset

Fig. 3. Student's perspective

3.1 Model description : Scenario

Consider an ordinary classroom with an intriguing technical twist: the teacher wears a special headset that enables him to see his students' holographic avatars, even while they are physically outside the classroom (Figure 4).

The Connected Teacher: Equipped with his holographic headset, teacher (Figure 2) prepares for a day of class. As he enters his traditional classroom, he activates his headset and sees a multitude of holographically teleported avatars of his students. The holographic representations of these characters are accurate; they are seated in their chairs and have a traditional table in front of them. Despite the physical distance, their motions and facial expressions are faithfully portrayed, giving the impression of a real presence. Note that holoportation requires more processing power and bandwidth, which is why we choose to communicate through avatars. Thanks to the spatial computing of the Hololens or the Meta Quest 3, our device offers a realistic scale and spatial awareness, with real-time video streams of the participants.

The Teleported Student: Meanwhile, on the other side of the globe, his students join from their respective homes, facilitated by our holographic telepresence device. While some are in Asia, others in Europe, they all convene virtually within the same classroom. Each avatar represents a real student, enabling real-time interaction with both the teacher and other students.Students simply connect with their computer and activate their webcam. Compared to Holoforge's solutions, this one is significantly lighter because no special camera is needed (Figure 3). This method makes remote learning more accessible to all students by enabling them to use easily accessible tools.

Fig. 4. Example of the proposed scenario

3.2 Implementation and demonstration

The implementation of our solution, as depicted in Figure 1, relies on several key steps, each contributing to the creation of an immersive and realistic experience. The first step involves detecting 3D positions using the MediaPipe library [4]. This library offers a robust suite of tools for real-time recognition and analysis of human movements. Pose estimation, MediaPipe's flagship feature, involves several steps:

Detection of key body points: MediaPipe identifies the main anatomical points of the human body, such as joints and extremities, from an image or video sequence.

Estimation of pose: Using these detected key points, MediaPipe reconstructs the human pose in 3D space, estimating the relative positions of different body parts (Figures 7 et 8). The second step in our process is to send the body's key points through a data transfer socket. This communication mechanism enables fast and efficient data transmission among different system components, thereby ensuring real-time synchronization between motion capture and processing. The final step in our implementation involves transcribing the detected points and lines into Unity, a platform renowned for game and simulation development. Using the data sent by the socket, we reconstruct the person's skeleton within Unity's virtual environment, enabling an immersive visualization of captured movements.

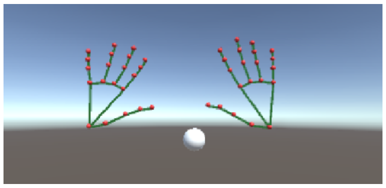

Fig. 5. Hand skeleton rendered in Unity using keypoints extracted from MediaPipe.

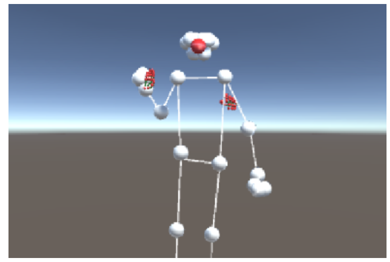

Fig. 6. Pose estimation of the body

After recreating the skeleton based on the body keypoints extracted by MediaPipe, we validated the accuracy and fluidity of the movements in Unity through real-time testing, demonstrating effective synchronization between motion capture and rendering. The next step is to introduce this skeleton into the Ready Player Me avatar[9]. Specifically, each detected keypoint (such as joints) will be mapped onto the rigging of the Ready Player Me avatar. This avatar, created from realistic photos, already has a mesh and a basic skeletal rig. We link the keypoints from the skeleton generated by MediaPipe to the corresponding bones of the avatar to ensure that the movements captured in real time are faithfully reproduced on the avatar. Once the skeleton is properly integrated into the avatar, we add this avatar to the 3D environment in Unity. In this virtual environment, the avatar is placed in an interactive setting where it can move and interact with the virtual elements around it. The Unity environment is designed to simulate collaborative spaces (such as a classroom), and the avatar, thanks to the real-time synchronization of movements, allows the user to fully immerse themselves in this space.

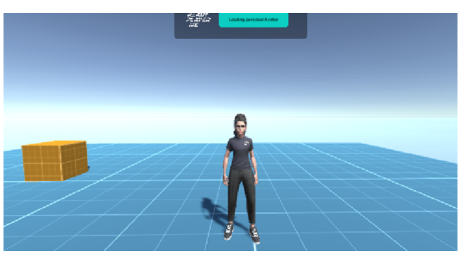

Fig. 7. ReadyPlayerMe avatar into the virtual environment

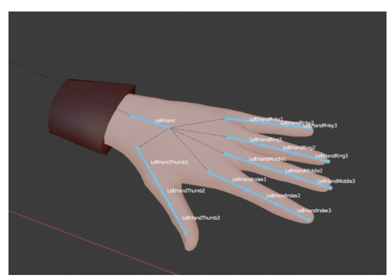

Fig. 8. Integration of key points in the avatar's hand

[9] Ready Player Me :`https://readyplayer.me/avatar`

4 Conclusion and Futur Works

In this paper, we present an innovative concept of mixed reality for hybrid education, combining pose estimation, holoportation, and mixed reality. Our goal is to develop an immersive experience that improves student involvement and collaboration in order to overcome the limitations of traditional distance learning methods.Further studies are needed to assess the pedagogical impact of our approach on student learning, comparing the performance of students using our system with those using traditional methods of distance teaching. In conclusion, our study paves the way for exciting new opportunities in the field of hybrid education and remote communication.

References

[1] Areej Banjar et al. "A systematic review of the experimental studies on the effectiveness of mixed reality in higher education between 2017 and 2021". In: *Computers Education: X Reality* 3 (2023), p. 100034. ISSN: 2949-6780. DOI: `https://doi.org/10.1016/j.cexr.2023.100034`. URL: `https://www.sciencedirect.com/science/article/pii/S2949678023000284`.

[2] Stephan Beck et al. "Immersive Group-to-Group Telepresence". In: *IEEE Transactions on Visualization and Computer Graphics* 19 (2013), pp. 616–625. URL: `https://api.semanticscholar.org/CorpusID:1031304`.

[3] C.E. Hughes et al. "Mixed reality in education, entertainment, and training". In: *IEEE Computer Graphics and Applications* 25.6 (2005), pp. 24–30. DOI: `10.1109/MCG.2005.139`.

[4] Indriani, Moh. Harris, and Ali Suryaperdana Agoes. "Applying Hand Gesture Recognition for User Guide Application Using MediaPipe". In: *Proceedings of the 2nd International Seminar of Science and Applied Technology (ISSAT 2021)*. Atlantis Press, 2021, pp. 101–108. ISBN: 978-94-6239-451-3. DOI: `10.2991/aer.k.211106.017`. URL: `https://doi.org/10.2991/aer.k.211106.017`.

[5] Sergio Orts et al. "Holoportation: Virtual 3D Teleportation in Real-time". In: Oct. 2016. DOI: `10.1145/2984511.2984517`.

[6] Charilaos Tsichouridis et al. "Virtual and Augmented Reality in Science Teaching and Learning". In: *The Impact of the 4th Industrial Revolution on Engineering Education*. Ed. by Michael E. Auer, Hanno Hortsch, and Panarit Sethakul. Cham: Springer International Publishing, 2020, pp. 193–205. ISBN: 978-3-030-40274-7.

Demo Description

During the demo session, we will present our holographic teleportation and mixed reality tool, specially designed for hybrid education. This innovative platform addresses the need for immersive, interactive, and engaging educational experiences in an era where remote and physical learning environments are increasingly blended. By leveraging advanced mixed reality technologies, we aim to show how teachers and students can transcend the limitations of physical space to create a truly collaborative classroom experience.

Conference attendees will witness how a teacher, equipped with a mixed reality headset (Hololens 2 or Meta Quest 3), interacts in real time with the holographic avatars of students, regardless of whether the students are physically present in the classroom or connected remotely. Our system is designed to enhance teacher-student interaction by bridging geographical divides, ensuring that all students, even those participating remotely, are "present" as life-like avatars. The teacher will perceive these avatars as though they were in the same physical room, which significantly improves the naturalness of communication, including the transmission of facial expressions, gestures, and other non-verbal cues, key to effective teaching.

The demo will feature a teacher and several students, all represented as 3D avatars within our mixed reality environment. MediaPipe's pose detection technology will be showcased as it captures the students' body movements in real time, which are then mapped onto their virtual avatars. This demonstration will highlight how the system accurately tracks motion and gesture, ensuring that the avatars' movements correspond to the students' real-world actions, providing an immersive and interactive experience that mimics real classroom dynamics.

We will provide an in-depth explanation of the underlying technical processes. Starting with real-time motion capture using MediaPipe's advanced pose detection system, we will demonstrate how the data is transmitted to our Unity-based platform. The 3D reconstruction pipeline, involving avatar creation, motion synchronization, and real-time rendering, will be explained step by step. We will also discuss the network architecture that enables seamless data transmission, ensuring minimal latency and providing a smooth experience for all users.

In addition to observing the interactions, conference participants will be invited to experience the system firsthand by creating their own personalized avatars using the "Ready Player Me" platform. By scanning a QR code provided on-site, attendees will be able to generate a 3D avatar based on real photographs and connect directly to the virtual environment. Once connected, participants will simulate a remote presence within the holographic classroom, observing how their avatar interacts with the teacher and other students in real time. This feature allows for an interactive and personal demonstration of our platform's capabilities.

Attendees will also have the opportunity to test the accuracy of gesture and facial expression synchronization by performing simple movements in front of their webcams. These movements, captured via a standard webcam, will be mirrored by their avatars in the virtual space, providing a real-time demonstration of the system's responsiveness. This interactivity not only enhances the user experience but also showcases the precision of the underlying pose estimation and facial tracking algorithms that drive our system.

While the demo highlights the capabilities of our holographic teleportation and mixed reality system, we will also discuss the current limitations and challenges associated with implementing such advanced technologies in hybrid education settings. One critical challenge is network connectivity, where issues such as latency, bandwidth requirements, and potential synchronization delays can affect the quality of the experience. We will explore these limitations in detail, discussing the technical hurdles and potential solutions. For instance, we will demonstrate how latency can cause delays in avatar motion rendering and explore future optimizations aimed at minimizing this effect.

Furthermore, we will present our plans for future developments, including enhancing avatar realism, improving gesture accuracy, and scaling the system for larger classrooms or educational institutions. We will also delve into how emerging technologies like 5G networks and edge computing could further optimize real-time holographic experiences, making them more accessible to a broader audience.

Our demonstration is not just a showcase of technology but a vision for the future of hybrid education. As we transition into an increasingly digital world, tools like ours could reshape the educational landscape by enabling teachers to engage students from anywhere in the world, fostering collaboration, and making learning more inclusive and interactive. The potential applications extend beyond traditional classrooms, enabling new forms of experiential learning in fields like medicine, engineering, and the arts, where hands-on interaction is critical.

In conclusion, this demo will provide a comprehensive look at the future of hybrid education through holographic teleportation and mixed reality. By enabling real-time interactions between teachers and students, regardless of physical distance, our system aims to redefine how we approach teaching in an increasingly globalized and digital world. We invite all attendees to participate in this immersive experience and envision how holographic technology could transform not only education but collaboration across industries.