HOLOVER-5G: Industrial Trainings Integrating Spatial Gaussian Splats with Real-Time 3D Holograms over 5G

Simon N.B. Gunkel[1], Tessa Klunder[1], Alessio Antoniacomi[2], Glauco Miniussi[2], and Gianluca Cernigliaro[1]

[1]*TNO,The Netherlands*
[2]*OVER, Estonia*

Corresponding author: Simon.Gunkel@tno.nl

Keywords: AR, MR, training and education at a distance, 3D representation

Abstract

In this (poster) paper, we outline our current work on HOLOVER-5G a scalable holographic communication pipeline deployed in 5G infrastructure and a commercial application. Our work integrates photo realistic 3D communication with environments as Gaussian splats to create high-quality training scenarios that are fully photo realistic and provide natural interaction.

Introduction

Next-generation eXtended Reality (XR), is enabling a revolution in human-centered applications, providing the possibility of enhancing shared digital experiences to become intuitive and realistic as never before. **Holographic communications** (Social XR) are increasing their level of realism and seamlessness, narrowing the gap between the virtual and the real, but encountering numerous challenges related to compression and transmission **with limited real-time capabilities** and **lack of scalable multi-user solutions**. Modern networking technologies such as 5G and 6G are increasing the confidence in the adoption and broad distribution of applications, such as volumetric video processing, especially when real-time is a requirement. HOLOVER-5G aims to provide the key technological contribution to reducing this gap, thanks to the introduction of the **Holographic Multipoint Control Unit** (**Holo-MCU)**, a virtualized optimization tool capable of i) offloading heavy processing from the client side; ii) maximizing the number of users in the remote experience; iii) providing elasticity to reconfigure and adapt the available network resources. Our work is based on / addresses the following 3 limitations.

Limitation 1, **Human presence**: Bringing real people in XR is key for realistic interactions. Remote XR experiences are normally based on virtual avatars that cannot adapt over time. Holograms provide this capability when dealing with natural content. With HOLOVER-5G, real humans remotely located will be in XR experiences. In particular in a specifically developed scenario for training in the manufacturing industry[9], allowing remote interaction, joint assets manipulation and a realistic feeling of being there.

Limitation 2, **Real-time**: holographic communications face challenges due to the large volumes of data (Viola et al., 2023) to be processed and solutions vulnerable to real-time constraints (Graziosi et al., 2020). HOLOVER-5G enables real-time communication between remote humans (Gunkel et al., 2023b) and advances the state of the art for live holographic communications with a network optimization solution (Gunkel et al., 2021).

Limitation 3, **Large-scale scalability**: Current pipelines address mostly peer-to-peer use cases (Jansen et al., 2022) without considering multiuser experiences, precluding the reaching of a bigger audience. We aim to overcome this lack applying multi-user optimization strategies to the volumetric video domain, exploiting 5G capabilities like stable mobile network connections and edge processing capabilities.

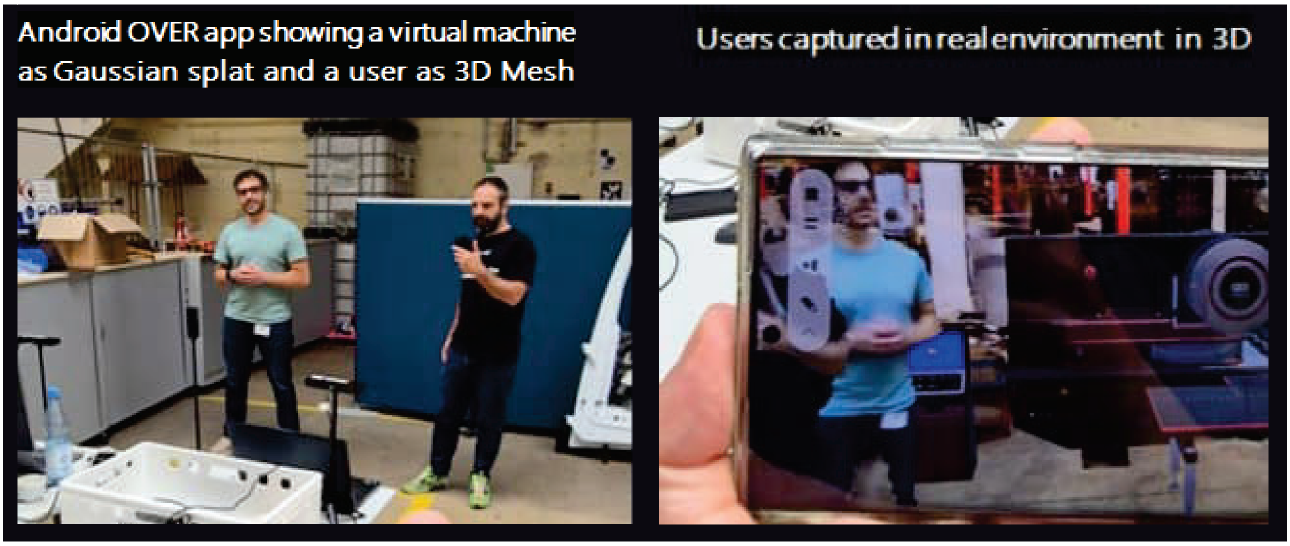

Figure P 1. *Initial demonstration of the use case in action (left: real environment, right: virtual mobile app)*

Use case

The main objective of HOLOVER-5G is to provide holographic training simulation within a highly realistic industrial environment (see Figure 1). At its core is a **Gaussian Splatting-based** Digital Twin of a factory, providing an **immersive** reconstruction of the facility. Within this environment, a 3D machine asset seamlessly transitions between multiple operational states, normal operation, error mode, and recovery, triggered by specific user actions. The **choice to include a static 3D asset with animated sequences** serves to further enliven the environment, making the demonstration **more engaging** and **immersive** for participants. By overlaying a holographic trainer onto the Digital Twin, users can observe machinery processes up close, practice guided troubleshooting, and refine their skills in a risk-free virtual setting that closely mirrors real-world conditions. Moreover, the scenario is generated with low complexity, allowing portability in multiple devices (i.e. smartphones or XR displays and XR glasses).

[9] https://doiotfieldlab.tudelftcampus.nl/ virtual-training-in-the-manufacturing-industry/

Technology

HOLOVER-5G proposes a complete end-to-end pipeline, as shown in Figure . We use a capture module as presented by Gunkel et al. (2021), per user, a singular RGBD (color and depth) capture sensor (ZED 2i) captures the user (and or objects) at color and depth, does some image optimizations (like cleaning and hole filling) and converts the 16bit depth values into our own greyscale color format (Gunkel et al., 2023a). Further, please note that the WebRTC output of the capture module will send our own modified VP9 encoding to enable the bitstream-based composition in the Holo-MCU (Gunkel et al., 2024).

At the core of HOLOVER-5G is a unity client integrated into a commercial application (OVER APP,[10]), the OVER APP leverages Gaussian splatting technology to deliver photorealistic, accurately scaled 3D reconstructions for industrial training scenarios. Each scene is pre-processed using an optimized pipeline, resulting in a lightweight, high-fidelity representation (under 30 MB per environment) that can be rendered in real-time on mobile devices. This approach enables immersive, interactive training experiences with realistic spatial cues, minimal storage requirements, and seamless performance, even on smartphones and tablets. Further, the client acts as a WebRTC Rendering Client that uses a custom shader to depict the RGBD points as custom 3D photo realistic mesh.

For the multi-user transmission, we use a Holo-MCU, based on the work presented in Gunkel et al. (2024) with the key difference of supporting WebRTC streams instead of RTP streams. The Holo-MCU connects with the individual upload clients via WebRTC to receive multiple VP9 RGBD streams. These streams follow a specific VP9 encoding (Gunkel et al., 2024). On each frame, the Holo-MCU combines all streams into one bitstream (in a horizontal tile composition). The final singular stream will be sent to all receiving user end clients.

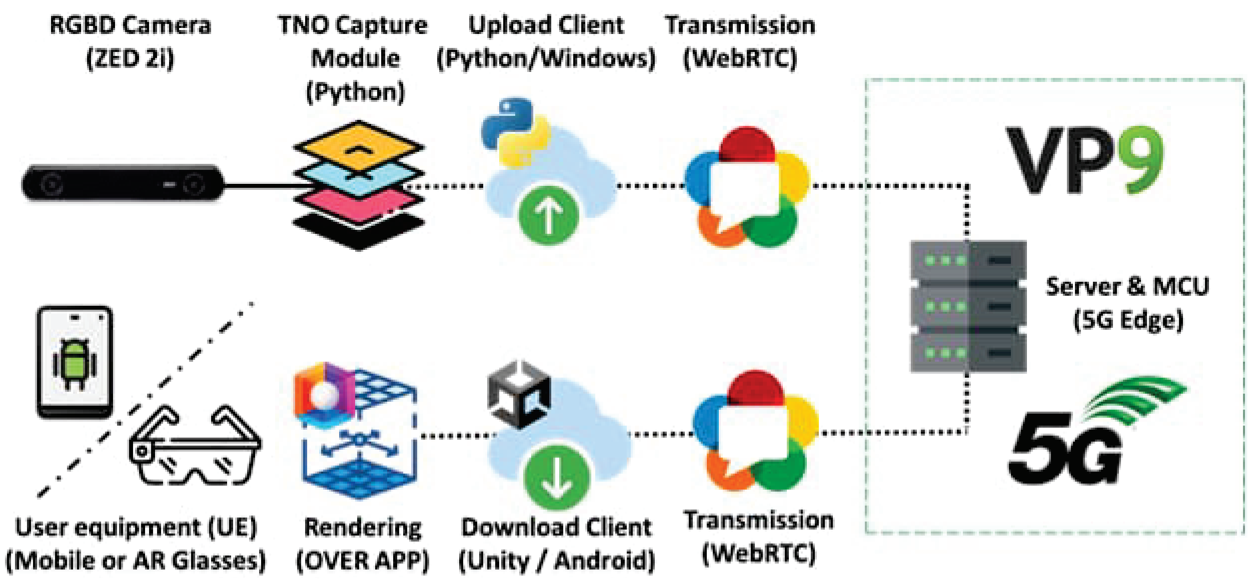

Figure P 2. Architecture and deployment.

[10] https://www.overthereality.ai/

5G deployment and initial results

As an initial evaluation of our technical approach, we deployed the Holo-MCU and end clients in a 5G Test Lab facility in Berlin. The Holo-MCU performance was tested locally running on a laptop (Acer Predator Helios 300, PH315-55s-917f) with pre-encoded RGBD streams (512x1024 pixel, 30fps) with 2, 4 and 8 tiles. System delay was tested with the DelAyrUco tool[11] with the Holo-MCU deployed in the testbed (edge) server and the DelAyrUco client connected over Wi-Fi. Thus, the setup is similar to Gunkel et al., 2024, but in a more realistic server deployment. Our measurements are shown in Table and indicate similar results as in previous experiments (Gunkel et al., 2024), but now in an operational environment. Important to note, however, is that we expect additional delays if end devices are connected via 5G.

Table P 1. Initial Measurements

Components	2 User	4 User	8 User
Holo-MCU CPU in %	1.77 (+/- 0.35)	2.57 (+/- 0.38)	3.46 (+/- 0.29)
Holo-MCU bitrate in MB/s	1.83 (+/- 0.47)	3.65 (+/- 1.28)	7.42 (+/- 2.09)
System Delay in ms	263 (+/- 30)	305 (+/- 35)	332 (+/- 36)

Conclusion and future work

In this paper we present our design and outline the technical of HOLOVER-5G to support enhanced technical trainings with the help of scalable holographic communication and Gaussian splats. Furthermore, we present initial measurement results evaluating the technical suitability of our approach. Finally, more technical and user tests are planned to test all aspects of our approach in a full 5G deployment.

Acknowledgements

This work was partially funded by the European Union: SPIRIT OC1 (101070672), ETHOS (Enhanced Training over Holographic Scalable Communications)[12].

References

Graziosi, D., Nakagami, O., Kuma, S., Zaghetto, A., Suzuki, T., & Tabatabai, A. 2020. An overview of ongoing point cloud compression standardization activities: Video-based (V-PCC) and geometry-based (G-PCC). APSIPA Transactions on Signal and Information Processing, 9, e13. https://doi.org/10.1017/ATSIP.2020.12

Gunkel, S.N., Dijkstra-Soudarissanane, S., Stokking, H.M., & Niamut, O.A. 2023a. From 2D to 3D video conferencing: Modular RGB-D capture and reconstruction for interactive natural user representations in immersive extended reality (XR) communication. *Frontiers in Signal Processing*, *3*, 1139897. https://doi.org/10.3389/frsip.2023.1139897

Gunkel, S.N., Klunder, T., & Cernigliaro, G. 2025. Ready Expert One: Universal 3D Workbench for Remote Industrial Training. In *Proceedings of the 2025 ACM International Conference on Interactive Media Experiences* (pp. 359–361). https://doi.org/10.1145/3706370.3731701

[11] https://github.com/TNO/DelAyrUco, 512x1024 pixel, 30fps
[12] https://spirit-project.eu/

Gunkel, S.N.B., Dijkstra-Soudarissanane, S., & Niamut, O. 2023b. "You AR' right in front of me": RGBD-based capture and rendering for remote training. In *Proceedings of the 14th Conference on ACM Multimedia Systems*. ence on ACM Multimedia Systems. https://doi.org/10.1145/3587819.3593936

Gunkel, S. N., Hindriks, R., Assal, K. M. E., Stokking, H. M., Dijkstra-Soudarissanane, S., Haar, F. T., & Niamut, O. 2021. VRComm: An end-to-end web system for real-time photorealistic social VR communication. In *Proceedings of the 12th ACM Multimedia Systems Conference*. https://doi.org/10.1145/3458305.3459595

Gunkel, S.N.B., Hindriks, R., Shiferaw, Y., Dijkstra-Soudarissanane, S., & Niamut, O. 2024. VP9 bitstream-based Tiled Multipoint Control Unit: Scaling simultaneous RGBD user streams in an immersive 3D communication system. In *Proceedings of the 15th ACM Multimedia Systems Conference* (pp. 23–33). https://doi.org/10.1145/3625468.3647608

Jansen, J., Subramanyam, S., Bouqueau, R., Cernigliaro, G., Cabré, M. M., Pérez, F., & Cesar, P. 2020. A pipeline for multiparty volumetric video conferencing: Transmission of point clouds over low latency DASH. In *Proceedings of the 11th ACM Multimedia Systems Conference*. https://doi.org/10.1145/3339825.3393578

Viola, I., & Cesar, P. 2023. Volumetric video streaming: Current approaches and implementations. In *Immersive Video Technologies*, 425–443. https://doi.org/10.1016/B978-0-32-391755-1.00021-3