Social eXtended Reality (XR) and Virtual Production: Crafting New Frontiers in Immersive Storytelling

Álvaro Egea Benavente
alvaro.egea@i2cat.net
i2CAT Foundation
Barcelona, Spain
University of Valencia
Valencia, Spain

Marc Martos Cabré
marc.martos@i2cat.net
MediaTech Area
i2CAT Foundation
Barcelona, Spain

Javier Montesa
jmontesa@brainstorm3d.com
Brainstorm Multimedia
Valencia, Spain

Francisco Ibañez
francisco@brainstorm3d.com
Brainstorm Multimedia
Valencia, Spain

Sergi Fernández
sergi.fernandez@i2cat.net
i2CAT Foundation
Barcelona, Spain

Mario Montagud Climent
mario.montagud@i2cat.net
i2CAT Foundation
Barcelona, Spain
University of Valencia
Valencia, Spain

Abstract

Immersive technologies are transforming how we consume, interact and engage with media. In this context, Social eXtended Reality (XR) is emerging as a powerful medium for remote communication, social interaction, and shared virtual experiences. Simultaneously, advances in multi-modal virtual production tools are expanding the boundaries of immersion, interactivity, and narrative design —while enabling the scalable distribution of rich immersive experiences via traditional 2D video platforms. This paper explores new frontiers in interactive and immersive storytelling through the convergence of Social XR, holographic communication, and virtual production. First, it proposes a cohesive architectural framework to seamlessly integrate innovative tools from these technological domains, aiming to support diverse interaction and media presentation modalities. Second, it outlines potential use cases and presents preliminary findings that demonstrate the feasibility and transformative potential of such newly envisioned cross-technology integration.

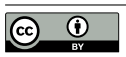
CCS Concepts

• **Information systems** → *Information systems applications.*

Keywords

Holographic Communications, Holoportation, Multimodal Interaction, Social eXtended Reality (XR), Storytelling, Virtual Production, Volumetric Video

How to cite this paper:
Álvaro Egea Benavente, Marc Martos Cabré, Javier Montesa, Francisco Ibañez, Sergi Fernández, and Mario Montagud Climent. 2025. Social eXtended Reality (XR) and Virtual Production: Crafting New Frontiers in Immersive Storytelling . In *Proceedings of ACM IMX Workshops, June 3 - 6, 2025*. SBC, Porto Alegre/RS, Brazil, 5 pages.
https://doi.org/10.5753/imxw.2025.1505



https://doi.org/10.5753/imxw.2025.1505

1 Introduction

Social Virtual or eXtended Reality (VR / XR) is emerging as a powerful medium for remote communication, social interaction and even collaboration across remote users. Previous studies have provided initial evidence on Social XR delivering enhanced quality of interaction and co-presence levels compared to traditional video-conferencing solutions [5], especially when supporting realistic and volumetric user representation formats instead of synthetic avatars [5][7].

In parallel, the availability of virtual production tools can help to coherently integrate heterogeneous (live and on-demand) and multi-modal content sources and interaction features in rich multimedia experiences, pushing the boundaries in interactive and immersive storytelling, as well as allowing a large-scale delivery of the resulting experiences via traditional 2D video distribution platforms [3][8]. A typical scenario in the TV broadcast ecosystem is that of a live presenter captured from a Chroma key room and teleported in real-time, with professional-grade stereo quality, to a virtual TV set where interactive 2D and 3D media content elements can be strategically orchestrated to provide advanced renders from the desired virtual camera position, overcoming intrinsic limitations and costs of physical setups.

This paper identifies and explores new frontiers in interactive and immersive storytelling through the convergence of Social XR, holographic communication, and virtual production. After briefly reviewing the state-of-the-art, the paper proposes a modular and cohesive architectural framework to seamlessly integrate innovative tools from these technological domains, aiming to support diverse rich media interactions and presentation features. Then, the paper outlines potential applicability scenarios and it reports preliminary findings from explored use cases that not only confirm the feasibility of the newly envisioned cross-technology integration but support its transformative potential and impact in the media landscape. Finally, the paper anticipates ideas and plans for future research.

2 Related Work

In recent years, both research and industry communities have devoted significant efforts to supporting social viewing and interaction scenarios, addressing not only traditional 2D formats but also emerging immersive video experiences [6].

In this context, the study in [4] provided initial evidence on the potential of (avatar-based) Social VR compared to traditional 2D conferencing tools, and to a baseline face-to-face scenario, for a multiuser photo sharing use case. Likewise, the study in [7] compared a shared video trailer watching experience, between couples, in three conditions: (i) face-to-face; (ii) avatar-based Social VR; and (iii) realistic video-based Social VR. It was found that the adoption of video-based Social VR resulted in increased levels of presence, togetherness and quality of interaction compared to the adoption of avatar-based Social VR tools, resembling closer baseline face-to-face experiences. The study in [5] provided further evidence on the potential of Social VR with realistic user representations, by: (i) integrating technological components to enable a real-time teleportation of mesh-based volumetric user representations to shared virtual environments; and (ii) creating and evaluating a 3D VR immersive story for which two teleported users need to gather and exchange hints to help understanding the whole narrative. In addition, the study in [2] shows that state-of-the-art technology for capturing and integrating realistic user representations into shared virtual environments, even when using single low-cost sensors for capturing the frontal users' viewpoint, already enables satisfactory and effective interactions and gesture-based collaborations when groups of two and four users meet virtually around a round table.

Simultaneously, virtual production techniques and tools are becoming increasingly adopted in the media industry, with a primary focus on broadcast but also addressing the filmmaking, creative industries, culture and even education sectors, given their multiple advantages in terms of flexibility, storytelling and cost-efficiency [3][8].

Finally, two recent pilot studies have provided initial – yet relevant and encouraging – evidence on the joint potential of Social XR, holographic communications and interactive virtual scenarios for the broadcast sector. First, the study in [1] showed the feasibility and promising receptiveness of Social VR for recreating interactive virtual TV shows, combining the real-time teleportation of remote participants represented as 3D holograms and a video-based presenter (billboard) from a Chroma key room, with an interactive and strategic presentation of multimodal content sequences to recreate a highly realistic and immersive experience. Second, the study in [6] reinforces the potential of such technologies to effectively recreate interactive time travel experiences, by allowing the teleportation of remote participants to a shared virtual environment where they can jointly explore a video catalogue, classified by year and content typology. With the selection of a particular year and video clip, the VR environment is adapted to resemble the associated epoque and content typology.

3 Integrated System Design

This section outlines the departing technological components, the identified requirements that motivate and drive the envisioned integration, and the resulting integrated architectural framework, including the adaptation and extension of the departing technologies with novel features and interfaces.

3.1 Departing technological components

Two core platforms / tools are proposed for integration:

- `HoloMIT`[1] `(Figure 1)`: Outstanding and versatile Social VR / XR tool, which allows a real-time volumetric and full-body capture and tele-transportation of remote users to VR/XR environments to socially interact, collaborate, and/or enjoy rich shared experiences (e.g., visits, performances) [5][1][2][6].
- `InfinitySet`[2] `(Figure 2)`: market-leading and award-winning virtual production tool designed for creating virtual sets, seamlessly integrating multimodal content and advanced interaction modalities in Virtual / Augmented / eXtended Reality (AR/VR/XR) scenarios. It acts as a powerful hub for integrating diverse technologies, like Chroma keyers, LED walls, tracking systems, and controllers, among others, perfectly integrated with current-practice broadcast workflows. InfinitySet also includes lightweight and user-friendly (cloud-based) releases targeted at other sectors beyond broadcast, like education and culture.

Figure 1: Departing technological solutions: HoloMIT

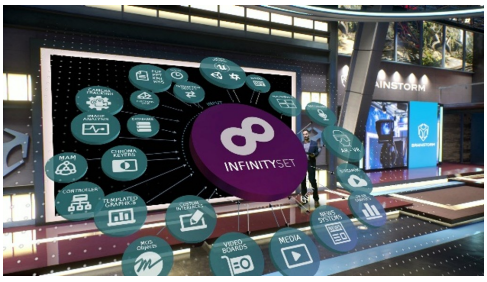


Figure 2: Departing technological solutions: InfinitySet

[1] https://i2cat.net/projects/holomit/ Last Accessed in May 2025
[2] https://www.brainstorm3d.com/products/infinityset/ Last Accessed in May 2025

3.2 3.2 Requirements for Interactive Storytelling

The strategic adaptation of, and integration between, both HoloMIT and InfinitySet solutions can provide new possibility for rich interactive storytelling. Diverse functional **Requirements (R)** have been identified in that regard, including:

- **R1.** Real-time integration of tele-ported 3D holograms, including the associated volumetric video and audio streams, into the virtual production engine.
- **R2.** Real-time integration of live presenters from a Chroma key room, including the associated stereo video and audio streams, into the virtual production engine.
- **R3.** Real-time distribution of the audiovisual streams from the live presenter(s) to the Social XR clients, as a new user representation format (beyond avatars and 3D holograms).
- **R4.** Support for seamlessly integrating and mixing video-based and synthetic content sources into the virtual production engine.
- **R5.** Dynamic generation and/or retrieval of content elements (e.g., 3D elements, Augmented Reality (AR) graphics, image / video sources…) into the virtual environment from the virtual production engine, and interactive distribution to the involved Social XR clients.
- **R6.** Support for loading and personalizing templates for virtual environments, including associated content elements and interactive presentations.
- **R7.** Translating interactions with the VR environment by the Social XR clients to the virtual production engine, so actions from remote participants can be perceived by the mass audience.
- **R8.** Coherent spatial rendering and layered content integration from specific positions and viewpoints in both the Social XR clients and virtual production engine.
- **R9.** Support for both immersive VR viewing and 2D viewing (e.g., so that the presenter can watch the other remote teleported users in the VR environment and the final produced program on flat screen).
- **R10.** Distribution of the produced program from the virtual production engine via traditional 2D video platforms, from any virtual camera with 6 Degrees of Freedom (6DoF).

3.3 Architectural Framework

Figure 3 sketches a high-level architecture of the resulting integrated framework combining the departing Social XR (augmented with holographic communications) platform and the virtual production engine, with the required interfaces, as well as adaptations and/or extensions to meet the target requirements.

On the one hand, the Social XR platform includes diverse client and server components to support shared multiuser communication sessions [1][2][6]. Regarding server components, a cloud Orchestrator is in charge of mainly session management and resources management, while two edge / cloud modules are in charge of audiovisual and interactions (e.g. events or control information) forwarding, following a Selective Forwarding Unit (SFU) strategy. The client components include modules for volumetric video capture, reconstruction and encoding, as well as for volumetric video decoding, composition and rendering (on 2D screens or on VR/XR headsets), along with their transmission and reception modules, respectively.

On the other hand, the virtual production engine mainly includes: (i) a module to receive an IP-based audiovisual stream conveying a professional-grade stereoscopic billboard of one or multiple users (e.g., the presenter/s), captured in real-time from a Chroma Key room, and including the associated audio stream; (ii) a wide set of multimodal content presentation and interaction features; and (iii) a module to render a 2D video stream and provide it to large-scale (broadcast or broadband) video distribution platforms.

The next key interfaces and extensions to such components have been devised to allow a seamless and robust integration between such components, meeting the target requirements:

- **Interface 1 (I_1).** It includes two evolved features: (i) Low-latency streaming channel to allow for successfully interaction with Social XR clients; (ii) sharing of camera tracking information from the presenter's capture setup, interfaced with the virtual production engine, to allow for a spatially coherent presentation of the associated content elements.
- **Interface 2 (I_2).** It includes diverse new and evolved features: (i) new Low-latency and high-capacity bi-directional data / streaming pipeline between a reference Social XR client (*Holo Client N+1* in Figure 3) and the virtual production engine, supporting content-agnostic data (e.g. audiovisual content, metadata like camera tracking data and interaction events…); (ii) rendering of the volumetric video content – as a modular layer – from a moving virtual camera inside the virtual environment of the reference Social XR client, and its delivery to the virtual production engine with high-resolution, high frame rate, and minimum latency; (iii) real-time distribution of the tracking-assisted presenter's billboard to all the involved Social XR clients, so that the presenter is perceived as an additional user (with a new representation type) for each remote Social XR participant; (iv) dynamic and interactive triggering of events and presentation of multimodal content elements from the virtual production engine to be distributed to all Social XR clients; (v) Interactive video and image (e.g., posters, slides) presentation from the virtual production engine to be presented in-sync to all Social XR clients; and (vi) reflection of interactions and content manipulation actions from Social XR clients to the virtual production engine, so these can be also perceived by the mass 2D audience.
- **Interface 3 (I_3).** It includes two main evolved features: (i) dynamic creation of virtual cameras from where to render the scene to be delivered to the mass 2D audience; (ii) remote control and positioning of the instantiated virtual cameras to deliver the experience from the desired positions and viewpoints.

4 Use Cases and Demonstrations

The resulting technology is planned to be evaluated and demonstrated for diverse relevant use cases and scenarios:

(1) *virtual TV debates / shows*, in which a presenter from a Chroma Key room is teleported in real-time to a virtual TV

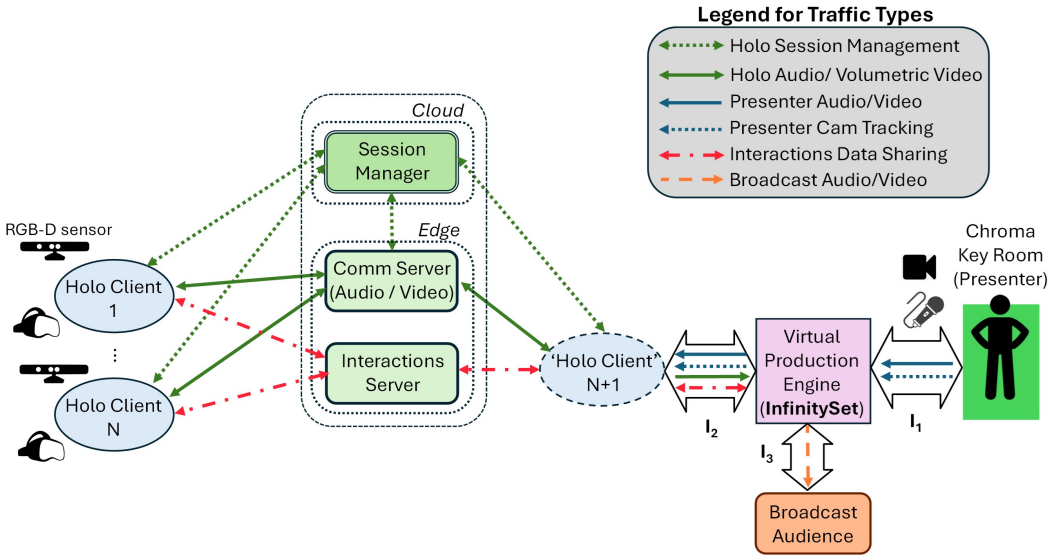


Figure 3: High-level architectural framework to integrate the Social XR, holographic communication and virtual production tools, with novel extensions

Figure 4: Social XR, holographic communication and virtual production tools strategically integrated to deliver rich interactive and immersive learning experiences, in the context of an anatomy lesson: (left) the trainer and a student when the parts of an arm are presented as a 3D model; (center) two students interacting with a 3D model of a lung; (right) the rendered 2D sequences from the 3D virtual environment, as a produced program, delivered back to the Social XR clients and presented on a virtual TV screen

and can interact with other remote teleported users represented as 3D holograms, while interactive content elements can be dynamically presented.

(2) *immersive learning and/or collaborative team building activities*, in which a instructor captured from a Chroma Key room can give a lecture or train remote participants in a virtual classroom or laboratory, with the possibility to interactively present and manipulate 3D models, 2D videos or slides, and even potentially to use interactive boards for brainstorming / co-design sessions. Figure 4 shows three captures from a demo of the resulting technology revolving around an anatomy lesson in a virtual environment.

All such experiences will be evaluated with both end and professional users, as done in previous works (e.g., [5] [6] [1] [2]), to get insights about its readiness, effectiveness, and to gather feedback on the awakened interest, potential, and extensions / optimizations to be applied to such a technology.

5 Conclusions and Future Work

This paper presented ongoing efforts to enhance immersive storytelling by strategically integrating Social XR, holographic communication, and virtual production tools. A fully functional architectural framework was developed, and its potential was illustrated through two representative media use cases.

Future research efforts will be concentrated on two main directions. On the one hand, the design, development, and integration activities will continue until all envisioned requirements and features are successfully provided, maximizing performance, immersion and co-presence aspects. On the other hand, the resulting technology will be assessed for diverse use cases, like virtual TV debates and education, which will serve to confirm its readiness and potential, as well as to identify aspects to be refined or newly provided.

Acknowledgments

This work has been partially funded by the European Union's Horizon Europe program, under agreement nº 101135637 (HEAT project), by the SNS-JU Horizon Europe Research and Innovation programme, under Grant Agreement 101096954 for 6G-BRICKS project, by Agencia Estatal de Investigación (AEI), in the framework of Proyecto Estratégico Orientado a la Transición Ecológica y a la Transición Digital (TED) 2021, under agreement TED2021-131690B-C32 and TED2021-131690A-C33 (REVOLUTION project), and by the Spanish Ministry of Economic Affairs and Digital Transformation and the European Union – NextGeneration EU, in the framework of the Recovery Plan, Transformation and Resilience (PRTR) (Call UNICO I+D 5G 2021, Ref. TSI-063000-2021-4 – 6G-OPENVERSO-HOLO). The work of Mario Montagud has been funded by MCIN/AEI/10.13039/501100011033 under Grant RYC2020-030679-I and by "the European Social Fund (ESF) Investing in Your Future".

References

[1] Sergi Fernández, Mario Montagud, Gianluca Cernigliaro, and David Rincón. 2021. Toward hyper-realistic and interactive social VR experiences in live TV scenarios. *IEEE Transactions on Broadcasting* 68, 1 (2021), 13–32.

[2] Sergi Fernández, Mario Montagud, Gianluca Cernigliaro, and David Rincón. 2022. Multiparty holomeetings: Toward a new era of low-cost volumetric holographic meetings in virtual reality. *IEEE ACCESS* 10 (2022), 81856–81876.

[3] Manolya Kavakli and Cinzia Cremona. 2022. The virtual production studio concept–an emerging game changer in filmmaking. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 29–37.

[4] Jie Li, Yiping Kong, Thomas Röggla, Francesca De Simone, Swamy Ananthanarayan, Huib De Ridder, Abdallah El Ali, and Pablo Cesar. 2019. Measuring and understanding photo sharing experiences in social virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.

[5] Mario Montagud, Jie Li, Gianluca Cernigliaro, Abdallah El Ali, Sergi Fernández, and Pablo Cesar. 2022. Towards socialVR: evaluating a novel technology for watching videos together. *Virtual Reality* 26, 4 (2022), 1593–1613.

[6] Mario Montagud, Marc Martos, Álvaro Egea, and Sergi Fernández. 2021. Toward hyper-realistic and interactive social VR experiences in live TV scenarios. *IEEE Transactions on Broadcasting* 68, 1 (2021), 13–32.

[7] Francesca De Simone, Jiel Li, Henrique Galvan Debarba, Abdallah El Ali, Simon Gunkel, and Pablo Cesar. 2019. Watching videos together in social virtual reality: An experimental study on user's QoE. In *2019 IEEE Conference on virtual reality and 3d user interfaces (VR)*. IEEE, 890–891.

[8] Jon Swords and Nina Willment. 2024. The emergence of virtual production–a research agenda. *Convergence* 30, 5 (2024), 1557–1574.