XRTC: Immersive Telepresence for First Aid and Emergency Scenarios

Pablo Perez[1,*,†], Ignacio Benito[1,†], Ester Gonzalez-Sosa[1,†], Marta Orduna[1,†] and Alvaro Villegas[1,†]

[1]*Nokia XR Lab, Madrid, Spain*

Abstract

We introduce XRTC, an immersive XR communication system that enables remote experts to obtain real-time spatial understanding of emergency scenes without requiring technical expertise from those on site. The system minimizes cognitive and operational burden for local non-experts, enabling precise guidance under stress.

Keywords

Extended Reality (XR), Immersive Communication, Remote Assistance, Emergency Response, XR for Challenging Environments

1. Introduction

Response time is critical when an individual experiences an accident or emergency, such as cardiac arrest. Although emergency services have developed protocols and equipment to maximize survival (such as emergency calls or publicly available defibrillators), current communication methods still rely largely on voice or video calls. In such situations, the person on site is typically not an expert and must describe the scene while simultaneously coping with stress. Communication barriers, contradictory information, incomplete descriptions, or unclear symptoms, can delay assessment and hinder the activation of appropriate emergency resources [1].

Nokia XR Lab has developed an immersive communication prototype (eXtended Reality Team Communication, XRTC) that allows a remote expert to virtually teleport into a physical environment in real time. This provides the remote user with rich spatial awareness of the scene, while enabling those on site to receive clear, actionable guidance without needing technical expertise or complex setup procedures. We propose applying XRTC to emergency situations, analogous to how automatic defibrillators empower non-professionals, enabling emergency services to obtain accurate, real-time situational awareness. This adaptation raises important technical, operational, and trust-related challenges. In this positioning paper, we outline the system's evolution, summarize its prior validation in related contexts, and discuss the implications of extending immersive telepresence to first-aid scenarios.

2. Extdended Reality Team Communication (XRTC)

2.1. What is XRTC

XRTC enables real-time immersive capture, transmission, and rendering of real environments, including the people within them. It is particularly suitable for hybrid scenarios involving both co-located (local) and remote participants who share a common physical space of interest. As illustrated in Fig. 1, a dedicated XRTC terminal (the **Snowl**), functionally similar to a large camera, is installed at a fixed location in the real world, typically surrounded by on-site participants (the **locals**). The terminal

CHI'26: XR4CE workshop, April 14, 2026, Barcelona, Spain

*Corresponding author.

†All authors contributed to the conceptualization, writing, and revision of this paper.

pablo.perez@nokia.com (P. Perez); ignacio.benito_frontelo@nokia.com (I. Benito); ester.gonzalez@nokia.com (E. Gonzalez-Sosa); marta.orduna@nokia.com (M. Orduna); alvaro.villegas@nokia.com (A. Villegas)

0000-0002-0999-0309 (A. Villegas)

Figure 1: XRTC applied to first aid scenarios. The upper part illustrates immersive emergency communication with a remote expert providing real-time guidance, while the lower part shows the evolution of the XRTC prototypes over time.

follows a plug-and-play design and becomes operational within seconds after power-on, an essential property in time-critical situations. Remote participants (the **travelers**) use a Virtual Teleportation application on a Head-Mounted Display (HMD), typically a VR headset. Upon launching the application, travelers are presented with an immediate sense of being co-located with the Snowl terminal. The travelers' experience is fully immersive (supporting 3-DoF head motion and 360 degree exploration), multimodal (integrating video and audio), and bidirectional. Travelers can communicate with locals and with other travelers, who are represented as avatars within the captured real scene, allowing multiple remote participants to simultaneously access and assess the same physical scene. The system can also render a self-body representation for the remote user, supporting embodied interaction and facilitating consultation of external resources during the session. At the physical Snowl location, locals see the travelers as avatars rendered on the terminal's integrated screen, and audio communication is provided via a built-in hands-free audio system. The teleportation illusion is strong, with users typically reporting a sustained sense of presence. The terminal's screen also serves as a Graphical User Interface (GUI) for configuring call parameters and device settings.

2.2. XRTC roadmap

The XRTC system has evolved through three main prototypes, each reflecting a shift from experimental validation toward a robust, deployable immersive communication solution. The first iteration, The Owl 1.0, was conceived as a proof-of-concept for laboratory testing, focusing on validating immersive telepresence using a simple setup augmented with an LED ring to convey remote gaze awareness [2]. This prototype enabled early use-case testing, but required on-site support and was not intended for standalone operation. The second iteration, The Owl 2.0, introduced a more compact configuration and incorporated commercial 360 degree cameras such as the Ricoh Theta and Vuze XR, enabling broader use-case experimentation beyond the lab [3]. This version supported XR Team Communication scenarios over private 5G networks and introduced companion applications: avatar-based representations of remote users on a tablet, and an smartphone application to use its camera as *magnifying glass* to show scene details to the remote user. While still requiring some operational support, this stage marked a transition toward more realistic deployment contexts. The current iteration, The Owl 3.0 (Snowl), adopts a fully integrated, plug-and-play design explicitly intended for unsupported testing and real-world deployment [4]. This version consolidates immersive capture, rendering, and interaction into a single self-contained terminal, prioritizing ease of installation, operational robustness, and reliability. As illustrated in Fig. 1, which presents the three prototypes at the bottom of the teaser image, this evolution highlights a clear progression from experimental systems toward a standalone XR communication device suitable for challenging environments such as emergency response.

2.3. Use cases and evaluation

The development of the XRTC solution has been done in parallel with many evaluation tests and activities, to assess its feasibility. The Owl prototype has been tested in contexts of professional and social meetings; primary, secondary and university education; remote expert support; psychological therapy. Structured experiments have been performed to assess several QoE features in conversation situations (using the 6 Thinking Hats technique [5]) or in exploratory ones, using a Treasure Hunt task [6]. This experiment has been a key asset to develop ITU-T Recommendation P.1321, which standardizes such methods. More recently, XRTC has also been used to support two use cases where a remote expert provided support to people with intellectual disabilities in two different contexts: supervision of daily tasks at home (telecare) and remote training of professional pastry operations [7]. This provides longitudinal evaluation of the system by the same group of people across 15 to 20 sessions over a 6-month period. Although these evaluations were not conducted in the emergency setting envisioned in this paper, they can be regarded as assessments of the XRTC user experience in analogous contexts, which we plan to extend to the emergency use case in future work.

Table 1 summarizes the results of the evaluations. Although results are not directly comparable, some high-level conclusions can be extracted: the sense of presence is very high, the audiovisual quality has significantly improved with the new version, and the cybersickness has been reduced similarly.

2.4. Adding AI-assisted augmentation

One of the key elements of the XRTC architecture is that it supports off-the-shelf integration with remote processing services. Based on this architecture, AI modules can be incorporated to provide real-time insights into the captured scene. In previous iterations of XRTC, this integration has enabled automatic augmentations of immersive communication, including detection of events of interest in remote-education classrooms [8], context-aware placement of remote avatars [9], and group-level affective understanding of participants [10]. These examples demonstrate that immersive telepresence can be systematically extended with backend intelligence without altering the user interaction paradigm.

From a systems perspective, such AI services may be deployed either in centralized infrastructures or at the network edge to reduce latency and improve data governance in time-critical scenarios [11]. This architectural flexibility is particularly relevant in emergency contexts, where near real-time analysis and privacy-aware processing are both required. Using the same integration model, more advanced AI systems can support scene interpretation and triage assistance in emergency scenarios [12]. Building on recent advances in agentic AI [13, 14], the XRTC architecture could further support an embodied AI co-pilot integrated directly into the immersive environment. Rather than functioning as a generic conversational agent, such a system would operate as a task-oriented cognitive teammate: retrieving protocols, structuring emergency procedures, monitoring adherence to guidelines, and providing decision support alongside the human expert. In this configuration, AI does not replace professional judgment but augments distributed expertise, acting as a stabilizing layer in high-stress situations. Given the critical nature of the scenario and the fact that the remote participant is an expert, the role of AI is to provide additional support that helps reduce the expert's cognitive load, for example by structuring scene information, monitoring adherence to protocols, or providing timely reminders. AI could also interact with the local user by offering additional information or visual guidance, but always under the supervision and coordination of the remote expert, and with explicit indication that the information or guidance is generated by an AI system.

3. XRTC as first aid companion: rationale and challenges

With these capabilities, XRTC can be deployed analogously to automatic defibrillators, establishing immediate communication with an emergency center and providing real-time situational awareness. Upon activation, the terminal would automatically initiate a call and transmit its GPS location, enabling rapid expert access. The remote expert, using an HMD, would obtain an immersive view of the scene to provide precise guidance and coordinate actions. When needed, multiple stakeholders, such as dispatchers or medical supervisors, could simultaneously access the same environment, supporting coordinated decision-making consistent with established models of interagency crisis collaboration [15]. Beyond audiovisual transmission, the system supports bidirectional metadata exchange, enabling integration of external sensors, scene annotations, or structured remote feedback. Mobile connectivity leverages 5G mechanisms such as traffic prioritization and secure authentication, contributing to robustness in safety-critical deployments.

When we compare the system to an automated external defibrillator, we refer specifically to the simplicity of use on the local side, where a non-expert should be able to power on or activate the terminal with minimal setup. Unlike an AED, however, XRTC does not operate autonomously in this context, but rather establishes a connection with an emergency professional. For the system to be viable, in addition to being robustly deployable across settings, emergency centers would need to have HMDs available for immediate use, and personnel would need to be trained in their operation.

3.1. Challenges

1. **Social, operational and ethical trust**. Adoption depends on acceptance by both local users and remote professionals. Trust must be calibrated to prevent over-reliance or under-utilization in safety-critical contexts [18]. This includes confidence in the communication channel, immersive representation, and AI-augmented insights, as well as addressing privacy, consent, and data governance in sensitive environments. In the current system, the communication setup has already been validated in remote support scenarios involving people with intellectual disabilities, suggesting that the user experience is suitable for non-expert users and that the immersive representation is effective for conveying scene information. However, the impact that AI assistance may have on local users, as well as the interaction between immersive technology and artificial intelligence, has not yet been validated.
2. **System robustness and resilience**. The architecture must tolerate connectivity fluctuations and prioritize emergency traffic over consumer use, ensuring continuity under stress. From a communications perspective, traffic prioritization mechanisms for emergencies are routinely used in communication systems for emergency and public safety services, and are already part of the network solution portfolios of operators and technology providers. It remains to be validated whether the XRTC communication system can be integrated with these prioritization mechanisms, and whether the system as a whole can maintain communication under degraded connectivity conditions.

Table 1
Key Performance Indicators (KPIs) across different test scenarios. MOS and cybersickness are in 1-5 scale (ACR and VSR respectively) [16]. Spatial and Social presence are in 1-7 scale [17].

Scenario	6 Thinking Hats [5]		Treasure Hunt [6]		Home Telecare [7]		Remote Training [7]
Owl Version	2.0		2.0		2.0		3.0
Magnifier	No		Yes		Yes		Yes
Remote Users	1		2		1		1
Local Users	2		2		3		1
KPIs	Remote	Local	Remote	Local	Remote	Local	Remote
QoE MOS	4.0	4.2	4.6	4.7	4.2	4.8*	4.8
Video MOS	-	N/A	3.1	N/A	3.6	N/A	4.3
Audio MOS	-	-	4.6	4.6	3.9	4.8*	4.1
Cybersickness (VSR)	-	-	3.9	N/A	4.1	N/A	4.9
Spatial Presence	4.7	N/A	5.9	N/A	6.8	N/A	6.9
Social Presence	5.2	5.2	5.8	5.8	6.7	6.7*	6.8

* These values are re-scaled from a 3-level scale adapted to people with intellectual disability

3. **Low-friction explainability**. Augmented information must remain immediately understandable under stress, requiring adaptive interfaces that present guidance or alerts without overwhelming users. In this respect, the evaluations carried out so far on adding supplementary information to XRTC remain limited, and need to be extended to the emergency use case and to operation under stress.

Acknowledgments

This research has been supported by the Eureka CELTIC-NEXT project C2025/1-10 6G Global Interslicing for Private Networks (6GLOBAL) funded by the Spanish Administration agency CDTI (EXP 00179761 / TIC-20250189).

Declaration on Generative AI

We used ChatGPT/Codex to assist with phrasing, translation from Spanish into English, and adaptation of the paper template; the ideas and scientific content are entirely the authors' own.

References

[1] V. Lindström, K. Heikkilä, K. Bohm, M. Castrèn, A.-C. Falk, Barriers and opportunities in assessing calls to emergency medical communication centre-a qualitative study, Scandinavian journal of trauma, resuscitation and emergency medicine 22 (2014) 61.

[2] R. Kachach, P. Perez, A. Villegas, E. Gonzalez-Sosa, Virtual tour: An immersive low cost telepresence system, in: 2020 IEEE conference on virtual reality and 3D user interfaces abstracts and workshops (VRW), IEEE, 2020, pp. 504–506.

[3] R. Kachach, S. Morcuende, D. Gonzalez-Morin, P. Perez-Garcia, E. Gonzalez-Sosa, F. Pereira, A. Villegas, The owl: Immersive telepresence communication for hybrid conferences, in: 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), IEEE, 2021, pp. 451–452.

[4] A. Villegas, E. Gonzalez-Sosa, P. Perez, J. Torres, The owl: Virtual teleportation through xr, in: 2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE, 2024, pp. 1158–1159.

[5] M. Orduna Cortillas, Understanding and Assessing Quality of Experience in Immersive Communications, Ph.D. thesis, Telecomunicacion, 2023.

[6] S. Baldoni, M. Orduna, A. Ferrarotti, M. Utiel-Moreno, K. Koniuch, P. Perez, F. Battisti, Methodologies for qoe evaluation in xr asymmetric collaborative applications, ACM Transactions on Multimedia Computing, Communications and Applications (2025).

[7] P. Pérez, D. González-Morín, J. Ruiz, J. Gutiérrez, M. Orduna, C. Cortés, M. Nava-Ruiz, M. J. Sánchez-Silva, Incluverso 5g: Designing and validating the technology for the incluverse, in: Workshop on Towards an Inclusive and Accessible Metaverse, ACM Conf. on Human Factors in Computing Systems, ACM, 2023, pp. 1–3.

[8] R. Kachach, M. Orduna, J. Rodríguez, P. Pérez, Á. Villegas, J. Cabrera, N. García, Immersive telepresence in remote education, in: Proceedings of the International Workshop on Immersive Mixed and Virtual Environment Systems (MMVE'21), 2021, pp. 21–24.

[9] A. Jiménez-Moreno, B. Rodriguez-Arbol, E. González-Sosa, P. Pérez, Á. Villegas, Context-aware remote user placement for immersive 360 communications, in: 2024 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR), IEEE, 2024, pp. 246–250.

[10] M. S. Martinez, E. Gonzalez-Sosa, J. T. Arjona, P. Perez, A. Villegas, M. Billinghurst, Vibe360: Group-based deep emotional understanding of immersive communications, in: 2025 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR), IEEE, 2025, pp. 261–265.

[11] W. Shi, J. Cao, Q. Zhang, Y. Li, L. Xu, Edge computing: Vision and challenges, IEEE internet of things journal 3 (2016) 637–646.
[12] A. Tahernejad, A. Sahebi, A. S. S. Abadi, M. Safari, Application of artificial intelligence in triage in emergencies and disasters: a systematic review, BMC Public Health 24 (2024) 3203.
[13] M. Abou Ali, F. Dornaika, J. Charafeddine, Agentic ai: a comprehensive survey of architectures, applications, and future directions, Artificial Intelligence Review 59 (2025) 11.
[14] L. Yan, From passive tool to socio-cognitive teammate: A conceptual framework for agentic ai in human-ai collaborative learning, arXiv preprint arXiv:2508.14825 (2025).
[15] N. Kapucu, Interagency communication networks during emergencies: Boundary spanners in multiagency coordination, The American review of public administration 36 (2006) 207–225.
[16] J. Gutierrez, P. Perez, M. Orduna, A. Singla, C. Cortes, P. Mazumdar, I. Viola, K. Brunnström, F. Battisti, N. Cieplińska, et al., Subjective evaluation of visual quality and simulator sickness of short 360 videos: Itu-t rec. p. 919, IEEE transactions on multimedia 24 (2021) 3087–3100.
[17] M. Orduna, P. Pérez, J. Gutiérrez, N. García, Methodology to assess quality, presence, empathy, attitude, and attention in 360-degree videos for immersive communications, IEEE Transactions on Affective Computing 14 (2022) 2375–2388.
[18] J. D. Lee, K. A. See, Trust in automation: Designing for appropriate reliance, Human factors 46 (2004) 50–80.