Exploring Privacy Challenges in Using Volumetric Video for Educational VR

Yu Liu
yliu2170@usc.edu
University of Southern California

Qiao Jin
qjin4@ncsu.edu
North Carolina State University

Feng Qian
fengqian@usc.edu
University of Southern California

Abstract

Volumetric video (VV) offers photorealistic 3D capture for immersive educational VR, often created by instructors through live-streamed lessons or prerecorded demonstrations. While enhancing engagement and presence, such instructor-produced content can unintentionally expose sensitive objects, personal information, or biometric identifiers, and may intensify feelings of surveillance. This poster examines these privacy risks in using VV for educational VR and presents a research agenda focused on integrating diminished reality (DR) techniques and real-time 3D scene understanding into VV pipelines to dynamically sanitize environments while balancing realism and privacy.

CCS Concepts

• **Applied computing → Interactive learning environments**; • **Security and privacy**;

Keywords

Volumetric Video, Privacy, Education, Virtual Reality

ACM Reference Format:
Yu Liu, Qiao Jin, and Feng Qian. 2025. Exploring Privacy Challenges in Using Volumetric Video for Educational VR. In *The Twenty-sixth International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing (MobiHoc '25), October 27–30, 2025, Houston, TX, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3704413.3765311

1 Introduction and Background

While video-based learning (VBL) has been widely used in online education, its integration into virtual reality (VR) environments, especially 360-degree videos, introduces new opportunities to enhance learner engagement by reducing environmental distractions and fostering a stronger sense of social presence. However, 360-degree video and other 2D formats in VR lack depth information and restrict learners to the originally recorded viewpoint, limiting spatial exploration and embodied learning [4]. Volumetric video (VV), also known as free-viewpoint video, addresses these limitations by capturing real-life subjects and environments in full 3D using multi-view photography or depth sensing [8]. A volumetric recording preserves the complete geometry and texture of the subject, including teacher's body pose, gait, microgestures, clothing, and accessories, alongside high fidelity details of the surrounding environment, such as office or lab layouts, bookshelves, whiteboard content, personal photos, and laboratory equipment. Viewers can also freely navigate inside volumetric scenes and extract perspectives that were never intended in the original capture. Compared with fully computer-generated VR environments, VV offers a more cost-effective production pipeline, as it does not require advanced 3D modeling or animation skills to achieve photorealistic scenes. Educational use cases already include virtual field trips, healthcare training, and creative storytelling [6], and show the pedagogical potential to enhance learners' engagement and interest in learning [3].

However, sharing pre-recorded or live-streaming VV content introduces unique privacy concerns compared to traditional video. Because volumetric capture reconstructs the entire surrounding environment in 3D, it can inadvertently include sensitive visual or audio information beyond the intended subject. For example, in a teacher's office, the capture might reveal private student records, personal photographs, intellectual property such as unreleased prototypes, or real-time contextual cues like people entering the background. The six degrees of freedom granted to the viewer mean that these details are not only visible from the original perspective but can also be discovered by exploring the scene from alternative angles or zoom levels after the recording. Once captured, such data can be stored, replayed, reprocessed, or even integrated into other virtual environments, raising the risk of secondary privacy breaches that persist far beyond the initial session.

One potential mitigation strategy is the use of diminished reality (DR) techniques in XR [1, 2], which selectively remove, replace, or obscure elements of the scene in real time before or during volumetric capture. These methods have been applied in VR telepresence and remote MR collaboration to protect sensitive data while maintaining spatial coherence for the viewer. However, applying DR in live VV streaming pipelines presents significant challenges, particularly in educational contexts: Captured scenes are often more complex, dynamic, and less tolerant to visual or semantic distortion due to the need for contextual learning [3]. For live-streaming scenarios, DR also introduces additional constraints, such as the need for real-time object detection, classification, and occlusion handling, all under strict latency requirements.

In this poster, we present an exploration of privacy issues in educational VV, with a focus on both pre-recorded live-streaming and live-streaming scenarios. We examine how the increased realism and presence of volumetric content influence both perceived benefits and privacy concerns and outline several research directions that will investigate teachers' and students' awareness of potential information leakage, their comfort levels with different capture settings, and the effectiveness of DR-based interventions in mitigating privacy risks.

MobiHoc '25, Houston, TX, USA

ACM ISBN 979-8-4007-1353-8/25/10
https://doi.org/10.1145/3704413.3765311

2 Privacy Concerns and Proposed Solutions

We discuss the challenges in both pre-recorded videos and live-streamed videos and present the potential issues and related research agenda aiming to resolve those challenges.

2.1 Pre-recorded Volumetric Video

Pre-recorded VV are applied in instructional lecture contexts [4]. Given the higher cost of capturing an entire class session, this poster focuses on pre-recorded videos featuring only the instructor. We outline the following privacy concerns and challenges associated with pre-recorded VV of instructors.

Sensitive Information Leakage. Volumetric videos capture the complete contextual environment surrounding the subject in a fully three-dimensional representation, thereby increasing the risk of exposing sensitive information within the instructor's environment, such as unintentionally visible personal documents, private items in the office or home, and whiteboard content containing unreleased research or student data. Compared to traditional 2D videos, volumetric formats also raise the cost and complexity of post-recording editing, making it more burdensome for instructors to identify and conceal privacy-sensitive objects.

Proposed Solution: We propose integrating 3D scene understanding models [7] to automatically identify privacy-sensitive objects and modify them directly within the video content, thereby preventing potential reverse-engineering or re-identification attacks. Ideally, such models should operate locally to eliminate the risk of data leakage through cloud processing.

Bio-information Inference Risk. As VV captures full-body motion in three dimensions, instructors are at risk of exposing biometric information such as facial models, body shape, gait, and gesture patterns, which can be used for recognition and cross-referencing across multiple recordings. An attacker can extract 3D facial models from VV and use them to impersonate a victim in deep-learning–based face authentication. Such biometric traits are hard to change once exposed, making their protection vital for preventing re-identification and preserving privacy.

Proposed Solution: We propose integrating privacy-preserving motion abstraction techniques that transform raw body geometry and kinematics into anonymized, semantically equivalent representations. Prior work has studied injecting imperceptible adversarial perturbations into volumetric streams [9] to avoid facial information detection. While effective at mitigating biometric leakage, such techniques may reduce the realism and quality of VV. Future research could focus on achieving an optimal balance between safeguarding privacy and preserving the social presence and engagement benefits of realistic volumetric content.

2.2 Live Volumetric Video

Live-streamed VV shares the privacy risks of pre-recorded content and adds challenges from real-time processing.

Real-time Processing and Error Sensitivity. During live streaming, unfiltered backgrounds, teaching spaces, or unexpected personal intrusions can be instantly broadcast, posing greater privacy risks than in pre-recorded videos, where content can be reviewed and edited before release. The real-time nature of streaming demands highly accurate and low-latency masking methods, as errors are far less tolerable—once privacy-sensitive content is exposed, it becomes immediately visible to the audience and may be recorded or redistributed without the instructor's consent.

Proposed Solution: While prior research has explored marking areas in the physical environment to automatically blur sensitive regions before broadcasting [2], this method cannot address unpredictable changes or newly introduced objects. We propose a hybrid privacy-preservation pipeline that combines lightweight real-time 3D scene understanding with proactive spatial zoning. The system would maintain a low-latency object detection and classification module for dynamic recognition, supported by a continuously updated "privacy buffer zone" that flags and masks new objects entering the scene. This approach aims to reduce computational overhead by focusing high-cost recognition only on changed areas, while providing immediate visual feedback to the instructor when masking is triggered, minimizing unnoticed privacy leakage.

Heightened Surveillance Pressure. Real-time capture and broadcasting can intensify perceived surveillance in online classes by making instructors and students feel continuously observed, evaluated, and recorded with limited control over context or audience. Prior studies of 2D live classes document how camera norms, monitoring interfaces, and policy patchworks increase feelings of being watched and constrain behavior [5]. Compared to 2D video, a fully 3D representation can amplify this effect by enabling multi-perspective viewing, persistent reinspection, and fine-grained behavioral analysis, further heightening surveillance pressure.

Proposed Solution: We suggest a formative study to examine psychological pressure experienced by instructors during real-time 3D streaming, using varied viewpoint counts, visual detail, and real-time analytics to identify key contributors to surveillance pressure.

References

[1] Yi Fei Cheng, Hang Yin, Yukang Yan, Jan Gugenheimer, and David Lindlbauer. 2022. Towards understanding diminished reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–16.

[2] Yuming Hu, Mingyu Zhu, Qiao Jin, Feng Qian, and Bin Li. 2023. MagicCloth: Protect User Privacy in AR Streaming. In *Proceedings of the 1st ACM Workshop on Mobile Immersive Computing, Networking, and Systems*. 222–228.

[3] Qiao Jin, Yu Liu, Yuxuan Huang, Bo Han, Feng Qian, and Svetlana Yarosh. 2025. An Exploratory Study of Using Interactive Volumetric Video in VR for Embodied Learning. In *Proceedings of the 19th International Conference of the Learning Sciences-ICLS 2025, pp. 2090-2094.* International Society of the Learning Sciences.

[4] Qiao Jin, Yu Liu, Puqi Zhou, Bo Han, Svetlana Yarosh, and Feng Qian. 2023. Volumivive: An authoring system for adding interactivity to volumetric video. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 569–570.

[5] H. Y. Lee, Y.-S. Kim, K. Kitani, and J. Wang. 2024. Investigating the Effects of Real-time Student Monitoring Interface on Instructors' Perceptions and Decisions. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery. doi:10.1145/3613904.3642845

[6] Yu Liu, Qiao Jin, Zejun Zhang, Bo Han, Svetlana Yarosh, and Feng Qian. 2024. HoloClass: Enhancing VR Classroom with Live Volumetric Video Streaming. In *Adjunct Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. 1–3.

[7] Songyou Peng, Kyle Genova, Chiyu "Max" Jiang, Andrea Tagliasacchi, Marc Pollefeys, and Thomas Funkhouser. 2023. OpenScene: 3D Scene Understanding With Open Vocabularies. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 815–824.

[8] Oliver Schreer, Ingo Feldmann, Sylvain Renault, Marcus Zepp, Markus Worchel, Peter Eisert, and Peter Kauff. 2019. Capture and 3D video processing of volumetric video. In *2019 IEEE International conference on image processing (ICIP)*. IEEE, 4310–4314.

[9] Zhongze Tang, Xianglong Feng, Yi Xie, Huy Phan, Tian Guo, Bo Yuan, and Sheng Wei. 2020. Vvsec: Securing volumetric video streaming via benign use of adversarial perturbation. In *Proceedings of the 28th ACM International Conference on Multimedia*. 3614–3623.