# Collaborative Scene Authoring with Near Real-Time 3D Reconstruction


Leon Foo
Centre for Immersification, Singapore Institute of Technology
Singapore, Singapore
leon.foo@immersification.org

Nirmal S. Nair
Centre for Immersification, Singapore Institute of Technology
Singapore, Singapore
nirmal@immersification.org

Liuziyi Liu
Centre for Immersification, Singapore Institute of Technology
Singapore, Singapore
ziyi@immersification.org

Jeannie S. Lee
Centre for Immersification, Singapore Institute of Technology
Singapore, Singapore
jeannie@immersification.org

Songjia Shen
Centre for Immersification, Singapore Institute of Technology
Singapore, Singapore
songjia@immersification.org

Indriyati Atmosukarto
Centre for Immersification, Singapore Institute of Technology
Singapore, Singapore
Indriyati@singaporetech.edu.sg

Alvin Chan
Singapore Institute of Technology
Singapore, Singapore
alvin.chan@singaporetech.edu.sg

Jing Shi
Singapore Institute of Technology
Singapore, Singapore
jing.shi@singaporetech.edu.sg

Yong Joo Loh
Tan Tock Seng Hospital
Singapore, Singapore
yong_joo_loh@ttsh.com.sg

Yih Yng Ng
Tan Tock Seng Hospital
Singapore, Singapore
yih_yng_ng@ttsh.com.sg

Michael Chia
Tan Tock Seng Hospital
Singapore, Singapore
alvariom@yahoo.com

Chek Tien Tan
Centre for Immersification, Singapore Institute of Technology
Singapore, Singapore
chek@immersification.org


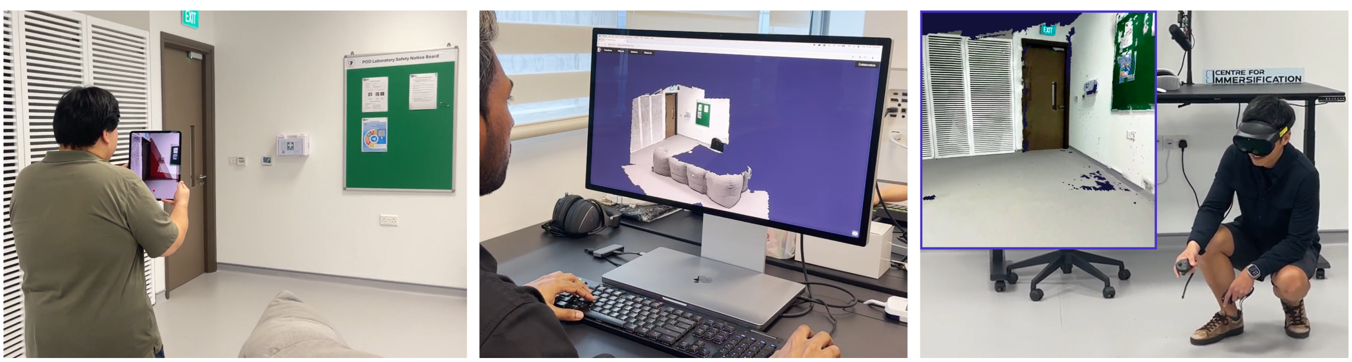


**Figure 1: A Capture Client user (left) actively scanning and reconstructing the environment, while an Editing Client user on a desktop (middle) and a VR user (right) navigate and interact with the digital reconstruction.**


## ABSTRACT

Traditional scene authoring workflows that demand realistic 3D environment scans often require a resource-intensive process involving data capture, processing, and model editing before integration into a 3D scene, typically using various hardware and software. This paper presents demoConstruct, an open-source collaborative platform that offers a unified 3D scene authoring environment. It enables simultaneous (1) 3D reconstruction and integration into a shared scene, (2) scene editing, and (3) VR exploration. During the interactivity session, attendees will experience the seamless capabilities of demoConstruct in collaborative capturing, editing, and experiencing of a shared 3D scene. This hands-on experience helps attendees explore innovative ways to leverage near real-time 3D reconstruction for advancing research in scene authoring, fostering the creation of new workflows and enhancing collaborative practices in the field.




*CHI EA '25, April 26-May 1, 2025, Yokohama, Japan*



ACM ISBN 979-8-4007-1395-8/2025/04.
https://doi.org/10.1145/3706599.3721190

## CCS CONCEPTS

• **Human-centered computing** → **Collaborative content creation**; **Computer supported cooperative work**; • **Applied computing** → *Computer-aided design.*

## KEYWORDS

3D Reconstruction, Collaborative Scene Authoring

**ACM Reference Format:**
Leon Foo, Nirmal S. Nair, Liuziyi Liu, Jeannie S. Lee, Songjia Shen, Indriyati Atmosukarto, Alvin Chan, Jing Shi, Yong Joo Loh, Yih Yng Ng, Michael Chia, and Chek Tien Tan. 2025. Collaborative Scene Authoring with Near Real-Time 3D Reconstruction. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3706599.3721190

## 1 INTRODUCTION

Collaborative scene authoring is a rapidly evolving field in Human-Computer Interaction (HCI) research, exploring diverse ways multiple users can work together to create, modify, and interact with 3D environments in real-time [2, 3, 15, 18].

For generating digital assets and environments, 3D reconstruction offers a cost-effective and efficient approach to digitizing real-world objects and scenes [4, 14]. In healthcare, this technology enables the digitization of medical equipment for training [11], while in the culinary industry, reconstructed dishes help diners make informed choices [1]. Traditionally, faithfully recreating real-world objects in digital form has required skilled artists, specialized software, and high-end equipment. By automating this process, 3D reconstruction offers a cost-effective and efficient alternative that lowers these barriers.

Recent advancements have significantly improved the accuracy and processing speed of 3D reconstruction technologies [8–10], making real-time applications increasingly viable [17]. However, practically accessing and utilizing these technologies in existing scene authoring workflows remains challenging due to their complexity, limited availability, and the technical expertise required.

The combination of collaborative scene authoring and 3D reconstruction is an emerging area with the potential to revolutionize how digital environments are created and modified [12, 16, 19, 20]. Yet, despite the growing body of research, to the best of our knowledge there is limited exploration on authoring environments that exploit near real-time 3D reconstruction algorithms to enable simultaneous reconstruction, editing and exploration tasks.

This paper describes demoConstruct, an open-source platform that enables accessible and collaborative scene authoring, powered by progressive 3D reconstruction (Fig. 1). Designed to accommodate both professionals and untrained users, the experience is delivered through web and mobile interfaces that run on common consumer devices. demoConstruct's progressive 3D reconstruction approach enables near real-time 3D reconstruction which supports simultaneous parallel operations across diverse user roles: (1) scene capturing and digitization, (2) scene editing, and (3) scene exploration. The version of demoConstruct presented in this interactivity builds upon our prior work, which established the system architecture [5] and showcased the technical features centered on progressive 3D reconstruction [6].

This interactivity session will offer attendees a hands-on experience with demoConstruct's collaborative features, allowing them to engage with its individual components by taking on user roles. Through direct interaction with demoConstruct, attendees will have the opportunity to reflect on and discuss how near real-time 3D reconstruction can be integrated into scene authoring scenarios, enabling novel workflows for their own use cases and projects.

## 2 DEMOCONSTRUCT SYSTEM

The demoConstruct system consists of two front-end clients: (1) the Capture Client, a mobile application for capturing reconstruction data, and (2) the Editing Client, a web-based application for editing and exploring reconstructions. These clients are supported by a back-end Edge Server, which handles data processing, 3D reconstruction, file storage, and API services. The progressive 3D reconstruction technique implemented in the Edge Server is adapted from the RTAB-Map SLAM-based reconstruction method [10]. More details on the system architecture and 3D reconstruction process can be found in our prior work [5, 6].

Both the Capture Client and Editing Client are designed for a streamlined user experience, featuring simplified tools and interfaces that reduce initial setup and minimize the learning curve. This approach ensures they are beginner-friendly, even for users with little to no experience in 3D reconstruction or scene authoring. Reconstructions generated by the Capture Client with the help of the Edge Server are automatically made available in the Editing Client, minimizing workflow friction for users, whether switching applications or collaborating with others in real-time.

The Capture Client is a mobile front-end application designed for consumer devices, enabling users to easily capture data for reconstruction, monitor the capture process, and view progressively reconstructed models (Fig. 2). To ensure a smooth data capture experience, spatial data is automatically saved every four seconds, processed, packaged into a database file, and uploaded to the Edge Server in the background, even as users continue capturing data. Upon receiving the database file, the Edge Server processes the data and generates a new incremental reconstruction version without requiring user intervention. Additionally, the Capture Client provides real-time feedback mechanisms to guide user actions during the capture process. An Augmented Reality (AR) overlay assists users by highlighting uncaptured areas with a red translucent layer, which disappears as those areas are captured. To help users assess the quality of reconstructions, the Capture Client retrieves the latest reconstruction version from the Edge Server, enabling immediate feedback. Users can evaluate whether the captured data is sufficient and if the reconstruction quality meets their requirements, providing guidance for their subsequent capture actions.

The Editing Client is a web-based front-end application that runs on any browser-supported device, such as desktop PCs or Virtual Reality (VR) headsets, without requiring installation. The Editing Client enables users to create shared scenes for exploring, inspecting, or manipulating reconstructions to build digital environments. Reconstructions are automatically scaled to real-world proportions, eliminating the need for manual resizing or adjustments. Like the

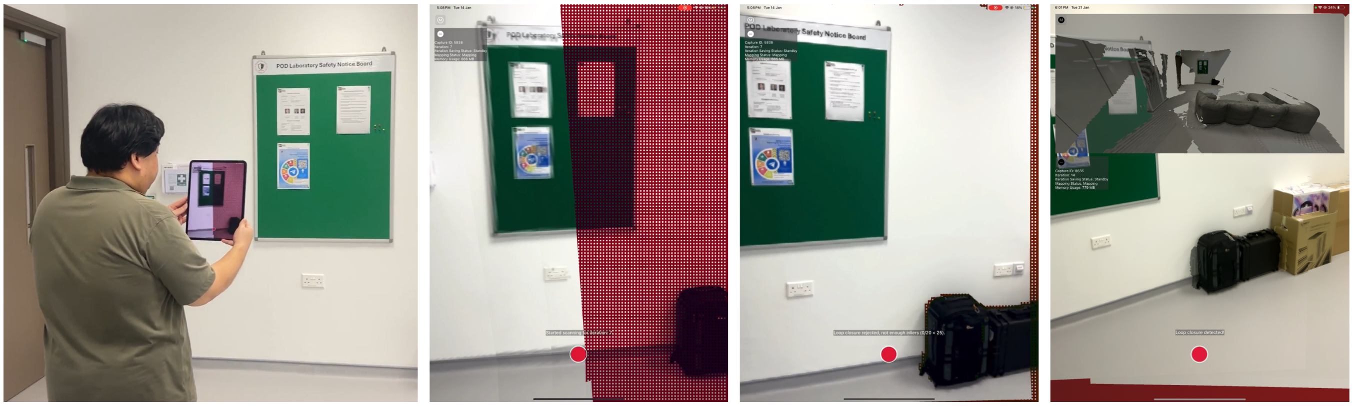

**Figure 2: An image sequence showing the scanning process from the Capture Client's perspective: (a) A user actively scanning an environment using a LiDAR-equipped iPad, (b) The Capture Client UI before areas are captured, indicated by a red translucent overlay, (c) The UI after the areas are scanned, where the red overlay disappears to indicate successful capture, and (d) An intermediate 3D reconstruction on the Capture Client, reflecting the progress of the scan.**

Capture Client, the Editing Client ensures reconstructions are continuously updated to their latest versions. This allows users to begin working with initial coarse reconstructions and seamlessly continue editing or exploring as finer versions become available.

The Editing Client supports multiple users interacting with the same shared scene in real-time. Depending on their roles, users can explore the reconstruction, collaboratively perform editing tasks, or both. By leveraging the spatial affordances of 3D reconstructions and a 3D workspace, users can perform tasks such as manipulating objects with transformation tools, adding metadata, overlaying annotations, and measuring object dimensions using the measurement tool (Fig. 3). These tasks, often challenging with static imagery, are made more intuitive with demoConstruct. When paired with a VR device, the Editing Client offers additional affordances through immersive interfaces. Users can perform natural, physical actions, such as using the measurement tool like a physical tape measure or navigating the VR play space to gain a grounded, spatial perspective. This immersive functionality enhances the intuitiveness and effectiveness of scene authoring and exploration workflows.

The Edge Server is the core back-end component of the demoConstruct system, housing the Reconstruction Module and serving as the central repository for storing capture data, 3D reconstructed models, texture files, and other associated assets. It also manages user data and shared scene information for both the Capture Client and Editing Client. The Reconstruction Module processes database files sent from the Capture Client to generate reconstructed models and their corresponding texture files. demoConstruct employs a coarse-to-fine approach for near real-time reconstruction, beginning with the rapid generation of lower-quality coarse models from the initial database files. As more capture data becomes available, the system processes larger database files to produce higher-quality fine reconstructions. This progressive approach allows for shorter processing times, typically just a few seconds for initial coarse models, enabling users to quickly view, edit, or assess reconstruction quality. As the capture process continues and additional data is gathered, previously generated coarse models are incrementally updated, giving users instant access to increasingly detailed and higher-fidelity reconstructions.

## 3 INTERACTIVITY EXPERIENCE

This interactivity session provides attendees with the opportunity to engage directly with demoConstruct, exploring how collaborative scene authoring can be applied to real-world scenarios and integrated into their own projects.

Participants will assume different roles and collaborate to capture, edit, and explore 3D scenes using demoConstruct's various clients (Fig. 1). Attendees will use the Capture Client to capture data for reconstruction, the desktop Editing Client to edit and manipulate the reconstructed scene, and the VR Editing Client to explore and inspect the scene in an immersive environment. Each participant's role will be determined based on their chosen device and client.

- **Scene Reconstructors** (Capture Client users): Reconstructors focus on capturing data of objects and environments for 3D reconstruction. Using the Capture Client, they can start, pause, and resume data capture as needed. Reconstructors can view the latest version of the reconstruction and additional details about the capture process, such as the number of reconstruction iterations completed.
- **Scene Editors** (desktop Editing Client users): Editors focus on modifying and enhancing the shared scene. Using the desktop Editing Client, they can place and reposition objects, measure distances, and annotate areas within the progressively reconstructed scene. These annotations can serve as guidance for themselves or other users (e.g., other Reconstructors and Inspectors) when placing additional objects or reconstructions.
- **Scene Inspectors** (VR Editing Client users): Inspectors explore the shared progressively reconstructed scene in an immersive VR environment. The VR Editing Client provides the same tools as the desktop Editing Client but adapts them

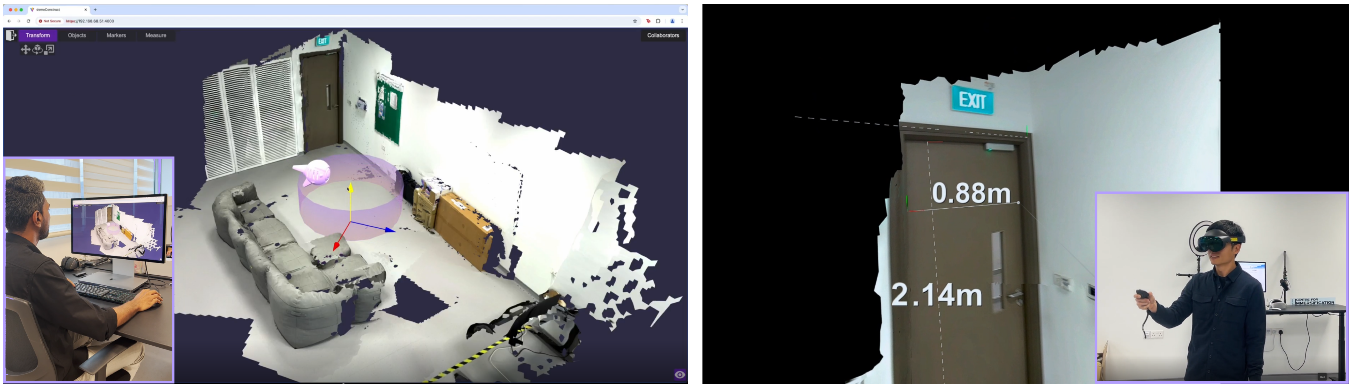


**Figure 3: An Editing Client user on a desktop places an object in the shared scene (left), while a VR user measures the dimensions of the door (right).**

for VR interactions. Measuring is accomplished through intuitive touch-and-select gestures, where users define a starting point and extend to an endpoint to calculate distances. Objects and markers can be placed and manipulated using point-and-click gestures. Locomotion within the VR scene is achieved through physical walking for a grounded experience, with teleportation as an alternative for navigating larger areas.

By demonstrating collaboration among these three user roles, each with distinct interaction modalities, demoConstruct highlights how near real-time 3D reconstruction and real-time cooperative workflows enable seamless collaborative 3D scene authoring. Tasks can be performed in parallel, reducing reliance on the completion of others' activities and streamlining the overall workflow.

In time-sensitive real-world scenarios, such as emergency disaster response situations, demoConstruct can function as a powerful real-time spatial planning and analysis tool. For example, ground rescue teams using the Capture Client can rapidly generate an initial coarse reconstruction of the environment, which is progressively updated in real time as the dynamic disaster scene evolves. Remote responders using the desktop Editing Client can immediately leverage the reconstruction to formulate plans without waiting for the extended time required to generate a fine reconstruction. At the same time, planners in the command centre can use the VR Editing Client to inspect the scene spatially in an immersive fashion and simulate actions on the plans created, enabling better coordination and faster decision-making.

In less urgent scenarios, demoConstruct can enhance processes by fostering co-design and collaboration. For instance, in rehabilitation therapy, it could support innovative remote consultation workflows. Caregivers could use the Capture Client to digitize a patient's living space, allowing remote therapists to review the conditions using the desktop Editing Client. Therapists could annotate the reconstruction or add objects to suggest modifications for improving the patient's living environment. Caregivers, in turn, could use the VR Editing Client to collaborate with the therapist in real-time, reviewing the suggested changes, proposing alternatives, or experiencing the recommendations from the patient's perspective, for example, by simulating navigation using a Personal Mobility Device (PMD).

## 4 CONCLUSION AND FUTURE WORK

This paper presents demoConstruct, an open-source platform for collaborative 3D scene authoring that integrates near real-time 3D reconstruction, scene editing, and VR exploration into a unified environment. By enabling multiple users with different interaction modalities to work in parallel on tasks such as data capture, scene editing, and scene inspection, demoConstruct streamlines collaborative workflows and enhances spatial understanding of real-world environments. Through its progressive 3D reconstruction approach, demoConstruct makes the process of scene authoring more accessible, reducing technical barriers and enabling real-time updates of digital content.

The hands-on interactivity session provides users with the opportunity to engage directly with demoConstruct's core features, experiencing firsthand how near real-time 3D reconstruction can enhance collaborative scene authoring processes. Through this session, attendees gain valuable insights into the platform's potential applications, demonstrating how demoConstruct can transform workflows across multiple industries, including disaster response and rehabilitation therapy.

Future work in development will focus on enhancing both the accessibility and performance of demoConstruct. This includes exploring more efficient methods for real-time 3D reconstruction to achieve faster processing [7] and higher visual accuracy [13]. These efforts are primarily aimed at reducing technical barriers for untrained users and improving the system's scalability to accommodate larger, more complex environments.

Future work in evaluation will involve conducting user studies to assess the experiential effects of dynamically updating virtual environments enabled by near real-time 3D reconstruction. Specifically, these studies include exploring ways to help users process and act on evolving, partial spatial information. The insights gained will inform future iterations of demoConstruct and contribute to the development of more efficient, accessible tools for collaborative 3D scene authoring across various industries and applications.

## ACKNOWLEDGMENTS

This research is supported by the National Research Foundation, Singapore, and Infocomm Media Development Authority under its Future Communications Research & Development Programme.

## REFERENCES


[1] Yashwith D Alva, Namrata Shrikant Janawade, Avani Jain, Nimish Bhasu, and Sarasvathi V. 2024. Augmented Reality Visualization of Menu Items using Image Reconstruction. In *Proceedings of the 2024 8th International Conference on Virtual and Augmented Reality Simulations* (Melbourne, Australia) *(ICVARS '24)*. Association for Computing Machinery, New York, NY, USA, 35–42. https://doi.org/10.1145/3657547.3657554

[2] Nicola Capece, Ugo Erra, Giuseppe Losasso, and Francesco D'Andria. 2019. Design and Implementation of a Web-Based Collaborative Authoring Tool for the Virtual Reality. In *2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*. 603–610. https://doi.org/10.1109/SITIS.2019.00123

[3] C.D. Cera, W.C. Regli, I. Braude, Y. Shapirstein, and C.V. Foster. 2002. A collaborative 3D environment for authoring design semantics. *IEEE Computer Graphics and Applications* 22, 3 (2002), 43–55. https://doi.org/10.1109/MCG.2002.999787

[4] Daniele Ferdani, Bruno Fanini, Maria Claudia Piccioli, Fabiana Carboni, and Paolo Vigliarolo. 2020. 3D reconstruction and validation of historical background for immersive VR applications and games: The case study of the Forum of Augustus in Rome. *Journal of Cultural Heritage* 43 (2020), 129–143. https://doi.org/10.1016/j.culher.2019.12.004

[5] Leon Foo, Songjia Shen, Ambrose Wee, and Chek Tien Tan. 2023. Progressive 3D Reconstruction for Collaborative Construction of Digital Twins. *Proceedings of the AAAI Symposium Series* 1, 1 (Oct. 2023), 7–10. https://doi.org/10.1609/aaaiss.v1i1.27466

[6] Leon Foo, Chek Tien Tan, Liuziyi Liu, Nirmal Sukumaran Nair, Songjia Shen, and Jeannie Lee. 2024. demoConstruct: Democratizing Scene Construction for Digital Twins through Progressive Reconstruction. In *ACM SIGGRAPH 2024 Labs* (Denver, CO, USA) *(SIGGRAPH '24)*. Association for Computing Machinery, New York, NY, USA, Article 2, 2 pages. https://doi.org/10.1145/3641236.3664424

[7] Huajian Huang, Longwei Li, Hui Cheng, and Sai-Kit Yeung. 2024. Photo-SLAM: Real-time Simultaneous Localization and Photorealistic Mapping for Monocular Stereo and RGB-D Cameras. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 21584–21593.

[8] Jungeon Kim, Hyomin Kim, Hyeonseo Nam, Jaesik Park, and Seungyong Lee. 2022. TextureMe: High-Quality Textured Scene Reconstruction in Real Time. *ACM Trans. Graph.* 41, 3, Article 24 (March 2022), 18 pages. https://doi.org/10.1145/3503926

[9] Lukas Koestler, Nan Yang, Niclas Zeller, and Daniel Cremers. 2022. TANDEM: Tracking and Dense Mapping in Real-time using Deep Multi-view Stereo. In *Proceedings of the 5th Conference on Robot Learning (Proceedings of Machine Learning Research, Vol. 164)*, Aleksandra Faust, David Hsu, and Gerhard Neumann (Eds.). PMLR, 34–45. https://proceedings.mlr.press/v164/koestler22a.html

[10] Mathieu Labbé and François Michaud. 2018. RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation. *Journal of Field Robotics* 36, 2 (Oct. 2018), 416–446. https://doi.org/10.1002/rob.21831

[11] Mingyang Liu, Dalei Guo, and Zhicheng Zhang. 2023. A Method of Touchable 3d Model Reconstruction based on Mixed Reality –A Case Study of Medical Training Applications. In *Proceedings of the 2023 6th International Conference on Image and Graphics Processing* (Chongqing, China) *(ICIGP '23)*. Association for Computing Machinery, New York, NY, USA, 83–89. https://doi.org/10.1145/3582649.3582679

[12] Hanuma Teja Maddali and Amanda Lazar. 2023. Understanding Context to Capture when Reconstructing Meaningful Spaces for Remote Instruction and Connecting in XR. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 275, 18 pages. https://doi.org/10.1145/3544548.3581243

[13] Hidenobu Matsuki, Riku Murai, Paul H.J. Kelly, and Andrew J. Davison. 2024. Gaussian Splatting SLAM. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 18039–18048.

[14] Ondrej Miksik, Vibhav Vineet, Morten Lidegaard, Ram Prasaath, Matthias Nießner, Stuart Golodetz, Stephen L. Hicks, Patrick Pérez, Shahram Izadi, and Philip H.S. Torr. 2015. The Semantic Paintbrush: Interactive 3D Mapping and Recognition in Large Outdoor Spaces. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3317–3326. https://doi.org/10.1145/2702123.2702222

[15] Michael Nebeling, Katy Lewis, Yu-Cheng Chang, Lihan Zhu, Michelle Chung, Piaoyang Wang, and Janet Nebeling. 2020. XRDirector: A Role-Based Collaborative Immersive Authoring System. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376637

[16] Benjamin Reynolds, James Hobin, Yuanzhi Cao, Martin Klaudiny, Daniel A Dangond, Kaize Xie, Keranie Theodosiou, Rik De Leeuw, Shiva Kashalkar, and Valentin Heun. 2024. The Pop-Up Metaverse: A Multi-User, Multi-Tasking Spatial Computing Environment for Collaborative Spatial Problem-Solving.. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '24)*. Association for Computing Machinery, New York, NY, USA, Article 424, 5 pages. https://doi.org/10.1145/3613905.3648677

[17] Keng Hua Sing and Wei Xie. 2016. Garden: A Mixed Reality Experience Combining Virtual Reality and 3D Reconstruction. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 180–183. https://doi.org/10.1145/2851581.2890370

[18] Theophilus Teo, Louise Lawrence, Gun A. Lee, Mark Billinghurst, and Matt Adcock. 2019. Mixed Reality Remote Collaboration Combining 360 Video and 3D Reconstruction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300431

[19] Kevin Yu, Ulrich Eck, Frieder Pankratz, Marc Lazarovici, Dirk Wilhelm, and Nassir Navab. 2022. Duplicated Reality for Co-located Augmented Reality Collaboration. *IEEE Transactions on Visualization and Computer Graphics* 28, 5 (2022), 2190–2200. https://doi.org/10.1109/TVCG.2022.3150520

[20] Domenic Zingsheim, Patrick Stotko, Stefan Krumpen, Michael Weinmann, and Reinhard Klein. 2021. Collaborative VR-Based 3D Labeling of Live-Captured Scenes by Remote Users. *IEEE Computer Graphics and Applications* 41, 4 (2021), 90–98. https://doi.org/10.1109/MCG.2021.3082267