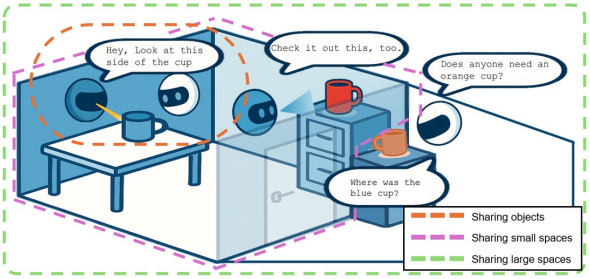
SpaceShare: Leveraging Multimodal Context for Fluid Sharing of Spaces in Video Meetings

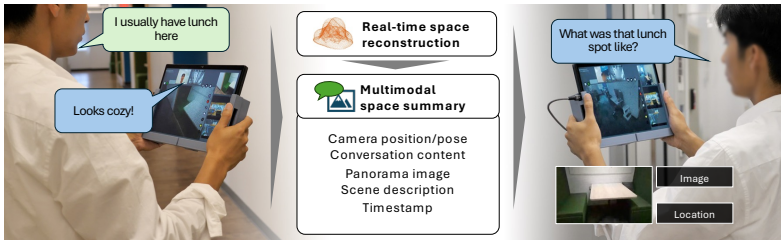
Author info removed for anonymity

Introduction



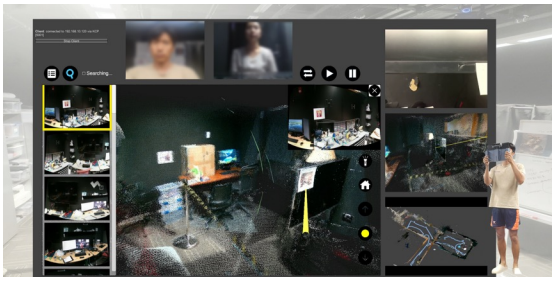
People often share objects and physical spaces during video meetings. Several projects have proposed solutions to make this sharing more efficient. However, sharing space remains challenging, especially in larger environments, due to the growing number of objects and locations, and the difficulty of navigating within the space.

Approach: Leveraging Multimodal Context



Meetings that take place in large spaces often involve numerous objects and locations, making recall difficult. While existing methods support search based on visual features or object types, these approaches can be insufficient when many similar items are present. To address this, we propose leveraging multimodal context, specifically, cues from user conversations, to augment the stored spatial data. This enables users to retrieve objects using more natural and relatable references.

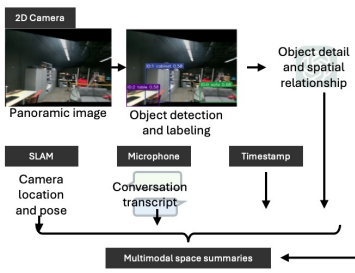
SpaceShare System



SpaceShare enables fluid spatial sharing in remote meetings by combining real-time 3D reconstruction with multimodal context. The system reconstructs the environment using SLAM, while capturing conversation transcripts, user positions, and object metadata. These are organized into spatial summaries, allowing users to retrieve specific scenes or objects using natural language.

3D Reconstruction

Multimodal Space Summarization



Context-based Space Retrieval