RealityCrafter: User-guided Editable 3D Scene Generation from a Single Image in Mixed Reality

Seokyoung Kim
seokyoung@kaist.ac.kr
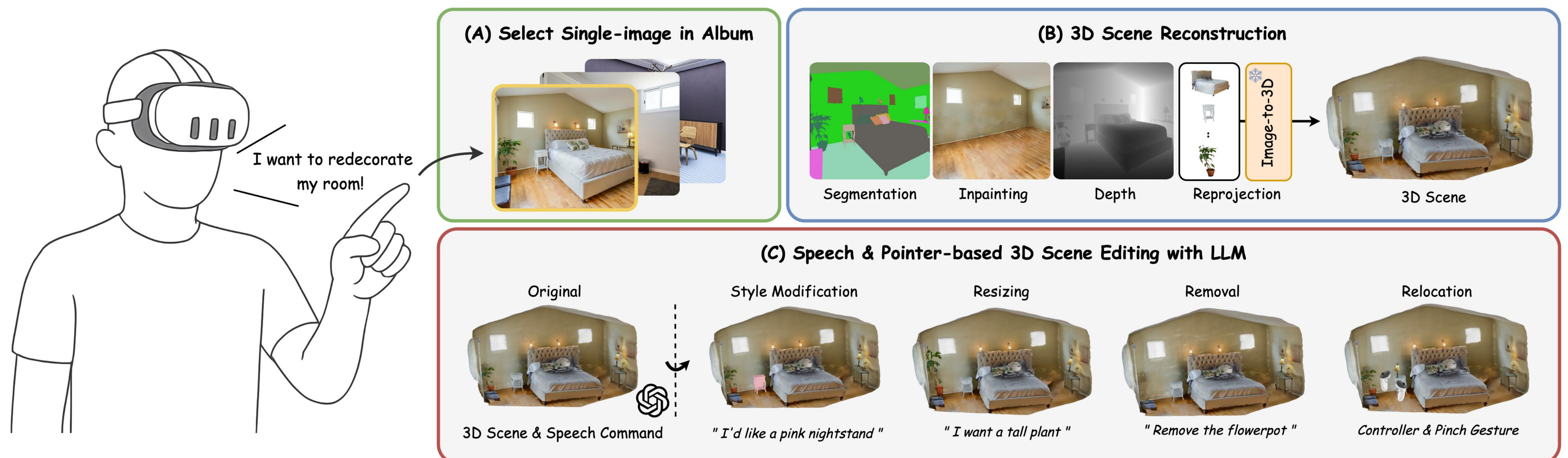
Dooyoung Kim
dooyoung.kim@kaist.ac.kr

Taejun Son
signal725@kaist.ac.kr

Woontack Woo
wwoo@kaist.ac.kr

UVR Lab, KAIST

(A) Select Single-image in Album

(B) 3D Scene Reconstruction

Segmentation — Inpainting — Depth — Reprojection — Image-to-3D — 3D Scene

(C) Speech & Pointer-based 3D Scene Editing with LLM

Original — Style Modification — Resizing — Removal — Relocation

3D Scene & Speech Command — " I'd like a pink nightstand " — " I want a tall plant " — " Remove the flowerpot " — Controller & Pinch Gesture

"I want to redecorate my room!"

RealityCrafter is a mixed-reality (MR) 3D authoring tool that **enables users to interact with and edit a 3D scene** reconstructed from a single real-world image.

Background

Problem:
Existing MR 3D authoring tools focus on purely virtual scenes or synthetic data, so they can't seamlessly blend and edit virtual objects within a user's real environment, breaking immersion.
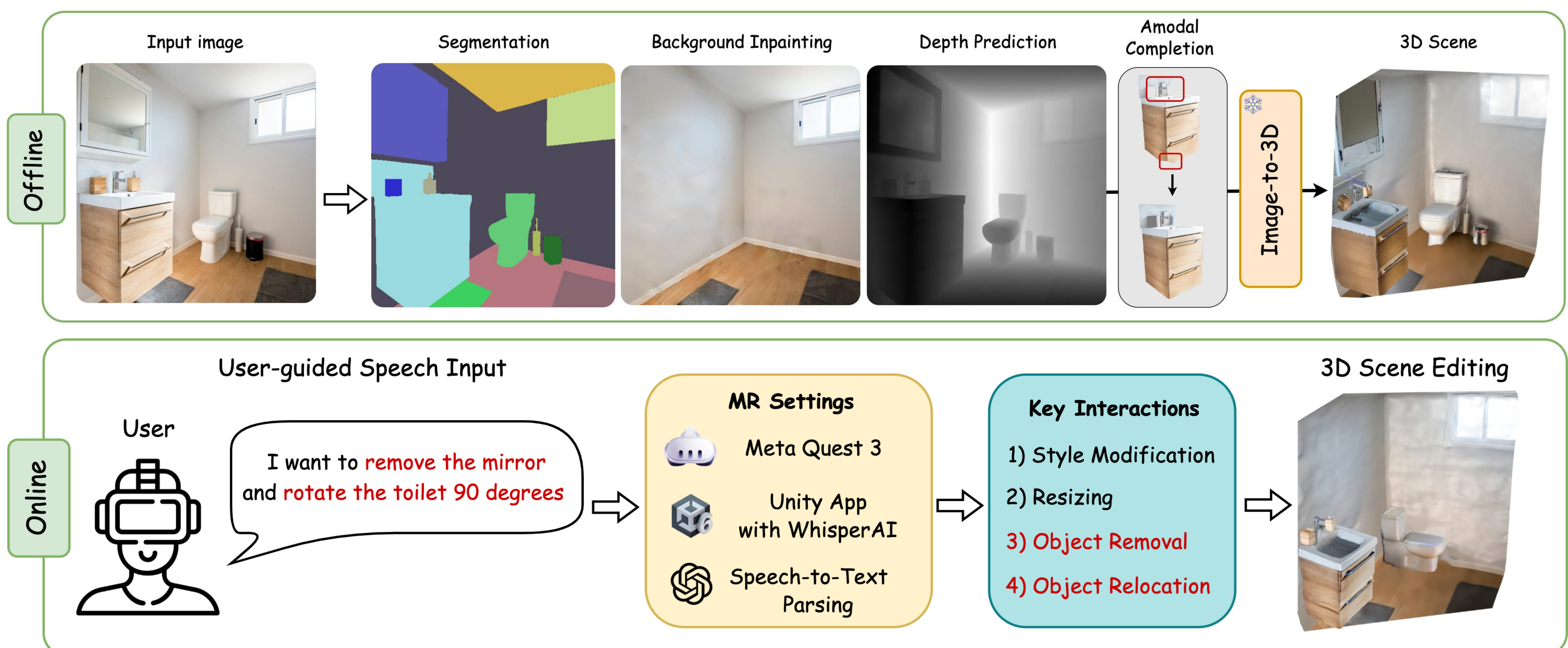
→ Proposed System:
We propose RealityCrafter, which allows users to experience and author spaces using only a single image of their actual environment, memories, or desired destinations.

System Implementation

The overall system consists of two main stages:
(1) offline; 3D scene reconstruction from a single image, and **(2) online; user-guided 3D object editing and authoring.**

Offline

Input image — Segmentation — Background Inpainting — Depth Prediction — Amodal Completion — Image-to-3D — 3D Scene

Online

User-guided Speech Input

User: "I want to remove the mirror and rotate the toilet 90 degrees"

MR Settings
- Meta Quest 3
- Unity App with WhisperAI
- Speech-to-Text Parsing

Key Interactions
1) Style Modification
2) Resizing
3) Object Removal
4) Object Relocation

3D Scene Editing

Conclusion

- We propose intuitive and immersive 3D authoring tool that enables users to edit and interacting with reconstructed 3D scene in mixed reality by immediately reflecting user intent through voice and pointer-based interactions.

- In future work, we plan to evaluate the system through user studies and validate it across real-world application scenarios, with the goal of extending it into a collaborative authoring tool for multi-user MR environments.

- We expect our system can revive spaces of memories, change present room interiors, letting users author 3D scenes.