Point Cloud Capture and Editing for AR Environmental Design

Zeyu Wang
Yale University
zeyu.wang@yale.edu

Cuong Nguyen
Adobe Research
cunguyen@adobe.com

Paul Asente
Adobe Research
asente@adobe.com

Julie Dorsey
Yale University
julie.dorsey@yale.edu

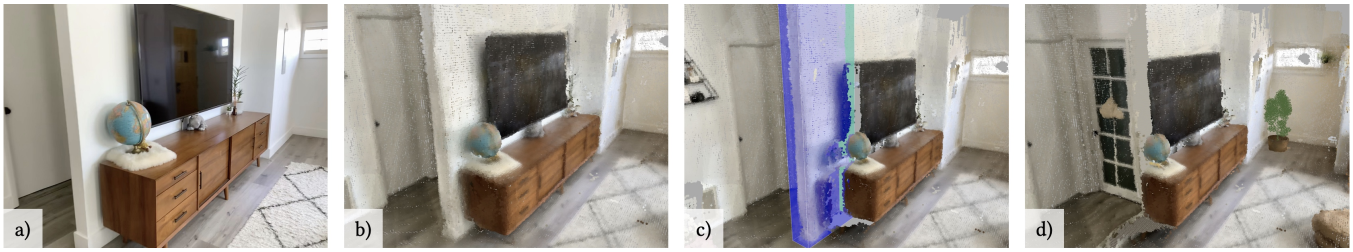

Figure 1: Overall workflow. a) Designers can quickly scan a physical environment like this living room on a consumer tablet. b) Our system constructs a localized point cloud in about 30 seconds. c) Designers can edit the scanned point cloud on the same tablet. d) The result of removing part of the wall, making it easier for residents to walk from the bedroom to the living room.

ABSTRACT

We present a tablet-based system for AR environmental design using point clouds. It integrates point cloud capture and editing in a single AR workflow to help users quickly prototype design ideas in their spatial context. We hypothesize that point clouds are well suited for prototyping, as they can be captured rapidly and then edited immediately on the capturing device in situ. Our system supports a variety of point cloud editing operations in AR, including selection, transformation, hole filling, drawing, and animation. This enables a wide range of design applications for objects, interior environments, buildings, and landscapes.

CCS CONCEPTS

• **Human-centered computing → Mixed / augmented reality**.

KEYWORDS

AR environmental design, point cloud, capture and editing

ACM Reference Format:
Zeyu Wang, Cuong Nguyen, Paul Asente, and Julie Dorsey. 2022. Point Cloud Capture and Editing for AR Environmental Design. In *The Adjunct Publication of the 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22 Adjunct), October 29-November 2, 2022, Bend, OR, USA.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3526114.3558636

1 INTRODUCTION

Environmental design involves visualizing changes to the physical environment and answering design questions [6]. For example, an architect might want to visualize how taking down a wall would affect the existing room arrangement. However, physical realization of a design can be costly if the design outcome turns out not to be desirable—a wall that was removed cannot be easily rebuilt. Modifying a replica of the environment is more common, but current practices like working with 2D blueprints or editing a model in 3D software require years of training to master.

Increasingly accessible consumer AR technologies promise a currently under-explored approach to environmental design. Previous AR scene editing systems require expensive hardware like HMD and camera array [2, 7], but 3D scene capture is much easier now with LiDAR-equipped iPads. Viewing the 3D scan in situ in AR is also supported out of the box. Users can quickly view the virtual 3D replica using natural movements and get an approximate sense of changes as if they had been made in the real space [5]. We present a tablet-based system that integrates point cloud capture and editing for environmental design. Our integrated workflow and the point cloud representation enable a faster feedback loop for designers to rapidly capture their spatial context, prototype environmental design ideas, and iterate their design through in-situ experiences.

2 SYSTEM OVERVIEW

Point cloud capture. We leverage LiDAR available on recent iPad Pro tablets to capture the 3D environment. Using ARKit, we record videos from the RGB and depth cameras. We also record estimated camera parameters and a depth-estimation confidence map for each frame. Our server processes the recorded data and recovers depth in meters. We build a truncated signed distance function (TSDF) volume [1, 4] and integrate the RGBD images with corresponding camera parameters. Finally, we can extract a colored point cloud from the TSDF volume for rendering and editing on the iPad.

Point selection. First, the user can tap on the screen to select an initial set of 3D points. Once they confirm the selection, our system automatically creates a bounding volume that covers the object of interest. The user can tap a face of the bounding volume and drag it for fine tuning. The user can also use touch-based interactions to move, rotate, and scale the bounding volume. Once the user has

UIST '22 Adjunct, October 29-November 2, 2022, Bend, OR, USA

ACM ISBN 978-1-4503-9321-8/22/10.
https://doi.org/10.1145/3526114.3558636

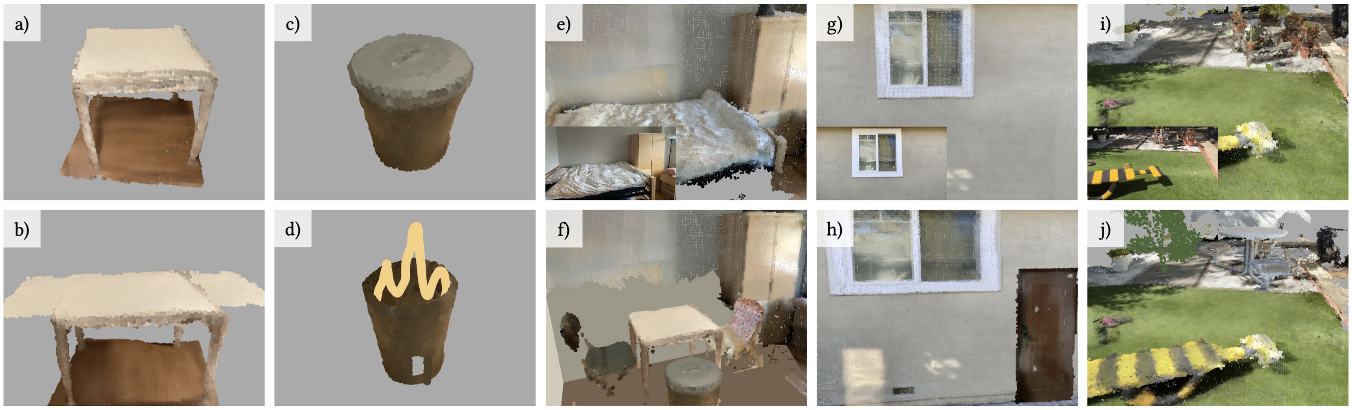

Figure 2: Applications. ab) Converting a square desk to a drop leaf table. cd) Repurposing a trash can as a stove. ef) Removing a bed from a bedroom and adding scanned furniture. gh) Making a window wider and adding a door. ij) Remodeling a backyard by adding a tree and a picnic table. Inset are reference photos in egi).

finalized the bounding volume, they have three choices: duplicate them to make a new point cloud object, remove them from the original point cloud object, or separate them from the original point cloud object—effectively duplication combined with removal.

Hole filling. Removing an object usually leaves a hole in the original scan. Holes may also exist because of incomplete scans. We developed an point cloud inpainting algorithm to fill the holes. The majority of holes are on one plane or are where two or three planes meet. The user can select a list of 3D points around the hole on each of the bounding planes. Our algorithm uses the singular value decomposition (SVD) to find a best-fit plane for each list of control points. We then compute the bounds and color for each fit plane and return a patch of points that the iPad adds to the scene.

Object manipulation. We implemented common touch-based interactions, such as a pan gesture for moving, a rotation gesture for rotation, and a pinch gesture for scaling. We added a control panel that can switch to different transformations, such as horizontal or vertical movement for pan, and scaling along one, two, or all three axes for pinch. To enable more freeform rotation in 3D, we leverage the iPad's pose tracking information and map any change in the iPad's pose to a corresponding change in the selected object's pose. This grabbing function manipulates objects in the physical scale, enabling intuitive operations like lifting a lid or opening a door.

Midair point drawing. The user can draw raw points in mid air to create more freeform structures by moving the iPad and pointing it in different directions. Our system creates and renders new points in real time, one meter in front of the camera.

Point cloud animation. We also provided an interface that lets the user create keyframes along a timeline. The user positions a timeline slider, makes changes to the scene, and adds a keyframe. Five attributes—position, rotation, scale, opacity, and point size—of each point cloud object are animated using linear interpolation.

3 APPLICATIONS

Object design. Designers can edit scanned objects in AR to add new functionalities or repurpose it for other use. For example, designers can convert a scanned square desk into a drop leaf table by first making the desk wider (Figure 2ab). They can then duplicate some points on the tabletop and add an animation to illustrate how the leaves are extended. Designers can use similar interactions to convert a scanned trash can into a stove (Figure 2cd), such as point selection and removal, 2D scaling, and midair drawing.

Interior design. Our system is also suitable for classic interior design tasks. A strength of our system is allowing designers to remove objects from their current environment similar to Diminished Reality [3]. It also allows designers to capture objects from other environments and import them to the current scan. For example, designers can convert a bedroom into a dining room by removing a bed and adding furniture scanned elsewhere (Figure 2ef).

Architectural design. Our system supports more freeform designs as it does not rely on predefined 3D assets or conversion to CAD models. This enables designers to prototype ideas for more open-end tasks like architecture design. For example, designers can modify the geometry of a building (Figure 2gh). They can select the window and make it wider, so more light and fresh air can come into the house. They can also add a door scanned elsewhere to see if it would be convenient to have a new entrance into the house.

Landscape design. Similar to architectural design, our system is also suitable for AR landscape design. For example, a designer can redesign a backyard after capturing it as a point cloud. They can remove some plants from a cluttered area of the backyard, making space to add a picnic table and a tree (Figure 2ij). The tree was converted from a mesh model and the picnic table was scanned at a hardware store. Designers can quickly edit these point clouds in situ and decide if they should return to purchase the picnic table.

In summary, our system provides a rapid prototyping experience for AR environmental design, enabled by integrating point cloud capture and editing on a single tablet device. This would have been very difficult to achieve physically, with a separate 3D scanner and a workstation, or with computationally expensive graphics programs for mesh processing and editing. We look forward to future research that makes scene capture and editing more accessible and provides a more intuitive AR prototyping experience using point clouds.

REFERENCES

[1] Brian Curless and Marc Levoy. 1996. A Volumetric Method for Building Complex Models from Range Images. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '96)*. Association for Computing Machinery, New York, NY, USA, 303–312. https://doi.org/10.1145/237170.237269

[2] David Lindlbauer and Andy D. Wilson. 2018. Remixed Reality: Manipulating Space and Time in Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173703

[3] Steve Mann and James Fung. 2002. EyeTap Devices for Augmented, Deliberately Diminished, or Otherwise Altered Visual Perception of Rigid Planar Patches of Real-World Scenes. *Presence: Teleoper. Virtual Environ.* 11, 2 (apr 2002), 158–175. https://doi.org/10.1162/1054746021470603

[4] Richard A. Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J. Davison, Pushmeet Kohi, Jamie Shotton, Steve Hodges, and Andrew Fitzgibbon. 2011. KinectFusion: Real-Time Dense Surface Mapping and Tracking. In *2011 10th IEEE International Symposium on Mixed and Augmented Reality*. 127–136. https://doi.org/10.1109/ISMAR.2011.6092378

[5] Zeyu Wang, Cuong Nguyen, Paul Asente, and Julie Dorsey. 2021. DistanciAR: Authoring Site-Specific Augmented Reality Experiences for Remote Environments. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 411, 12 pages. https://doi.org/10.1145/3411764.3445552

[6] Wikipedia. 2022. Environmental Design. https://en.wikipedia.org/wiki/Environmental_design. Accessed Jul 19, 2022.

[7] Ya-Ting Yue, Yong-Liang Yang, Gang Ren, and Wenping Wang. 2017. SceneCtrl: Mixed Reality Enhancement via Efficient Scene Editing. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. Association for Computing Machinery, New York, NY, USA, 427–436. https://doi.org/10.1145/3126594.3126601