Augmented Reality indoor navigation with x-ray visualization

by

© Francois Low

A Project Report submitted to the

Department of Computer Science

in partial fulfillment of the

requirements for the degree of

Bachelor of Science (Honours)

Supervisors: Dr. Matthew Hamilton, Dr Oscar Meruvia-Pastor

Department of Computer Science

Memorial University of Newfoundland

December 2022

St. John's Newfoundland

Abstract

A key problem in indoor navigation is spatial awareness and wayfinding. Augmented Reality systems can provide useful information and reduce perception load especially in ambiguous navigation situations. One of the challenges of indoor navigation is to locate a section of the building that is occluded by multiple layers of the interior structure of the building in the least time possible. In this paper, we implement an indoor navigation system on Microsoft HoloLens 2 that is capable of visualizing occluded infrastructure directly in the user's view of environment, thus enabling the users to see through the walls. In order to achieve this goal, we approach the problem with spatial mapping and x-ray visualization technique. The implemented system contains two functionalities: Mapping and Navigation. The mapping part utilizes Azure Spatial Anchors (ASA) to create a path from the origin to the destination and also 3D reconstruction of the landmark into 3D mesh model using Microsoft MRTK 2 SpatialMapping API. During navigation, a pathway of spatial anchors and a 3D mesh model of the landmark will be rendered in order to guide the user during navigation.

Acknowledgments

I would like to express my deepest gratitude to the following people, without whom I would not be able to complete my honours project. Dr Matthew Hamilton and Dr Oscar Meruvia-Pastor for their expertise, resources and their patience.

Contents

List of Figures

Chapter 1

Introduction

Traditional and digital 2D maps such as Google Maps are currently used as a way for outdoor navigation and route planning [1]. Outdoor navigation typically uses GPS for localization and path-finding [1]. 2D maps provide a topographic view which is useful for outdoor navigation. However, for indoor navigation, topographical view of the building is not effective and intuitive for users to navigate around a building especially if the destination is occluded another structure in the building. Complex building structures and convoluted floor plans are often unintuitive to first-time visitors to follow. While this might be a minor inconvenience to guests of a shopping mall, spatial awareness and wayfinding are mission critical to first responders like firefighters to respond and locate the affected part of the building in a limited amount of time. Furthermore, for the military, reconnaissance of urban terrain is often a critical step before offensive operations are conducted [2]. Knowledge of the layout of the hostile environment will give the offensive team an upper hand in offensive maneuvers. Previous works in AR indoor navigation involves using virtual visual cues such as arrows, transition lines or textual annotations to aid navigation. In this paper, we present an approach to implement a navigation system that can visualize occluded infrastructure directly in a user's view of the environment, allowing the user to explore hidden structure within the real-world environment.

Chapter 2

Related Works

2.1 3D mapping and localization for indoor environment

Mapping of indoor environment is a fundamental process for indoor navigation system. Generation of a 3D model of a landmark require the use of sensors (camera, depth sensor, lasers, LiDAR, etc) to collect indoor scene data. [3] presents a survey of 3D reconstruction, spatial mapping and localization of indoor scenes using Microsoft HoloLens. Although Microsoft HoloLens was not originally designed to be a indoor mapping device, results from the paper shows that Microsoft HoloLens is able to produce high quality reconstruction of 3D mesh models and localization of indoor scenes. Similar works like [4] utilizes ArUco markers for localization of indoor structures in order to visualize BIM (Building Information Modeling) data with Microsoft

HoloLens.

2.2 Visualization techniques in augmented reality

Visualization, the process of converting abstract data into visual representation [5], has played a big role in augmented reality applications. Visualization techniques in AR aim to reduce information clutter, direct attention and provide depth perception with the use of visual cues. [5]. Various visualization techniques have been explored in previous works to tackle problems faced in integrating digital contents into our view of the real world. One of the key problems in AR is the ability to determine the depth order between the real world and the digital objects. One use case would be overlaying digital objects that are occluded by multiple layers of physical structures in the real world while providing depth information to the viewer. [6] presented an overview of the ghosting pipeline that utilize the blending of physical depth cues to produce a convincing integration of occluded virtual objects into the physical world. The main idea of ghosting is to preserved selected features (edges, salient regions and texture details) of the physical world and render it on top of the digital objects [6]. [7] approach this problem using appearance cues to provide depth estimation. In this paper, the authors propose the use of display attributes (opacity and intensity) of 3D model renderings to encode depth of occluded buildings.

2.3 Augmented Reality Navigation Systems

Several studies have shown that augmented reality navigation systems provided virtual cues that can enhances users' surrounding awareness and reduces perception load. [8] presents an AR-bassed indoor navigation system that utilizes RGB-D camera and Oriented Brief Simultaneous localization and mapping (ORB-SLAM) algorithm to generate a point cloud of the environment and generate a hybrid map that used to provide physical visual cues by visualizing the landmarks' names on the 3D space.

Chapter 3

Methodology

3.1 Data

3.1.1 Study Area

The study area is restricted to one building. The Engineering building (ENGR) of Memorial University is chosen as the study area. For the demonstration of this project, spatial mapping and navigation are conducted on the Level 2 and 3 of the engineering building.

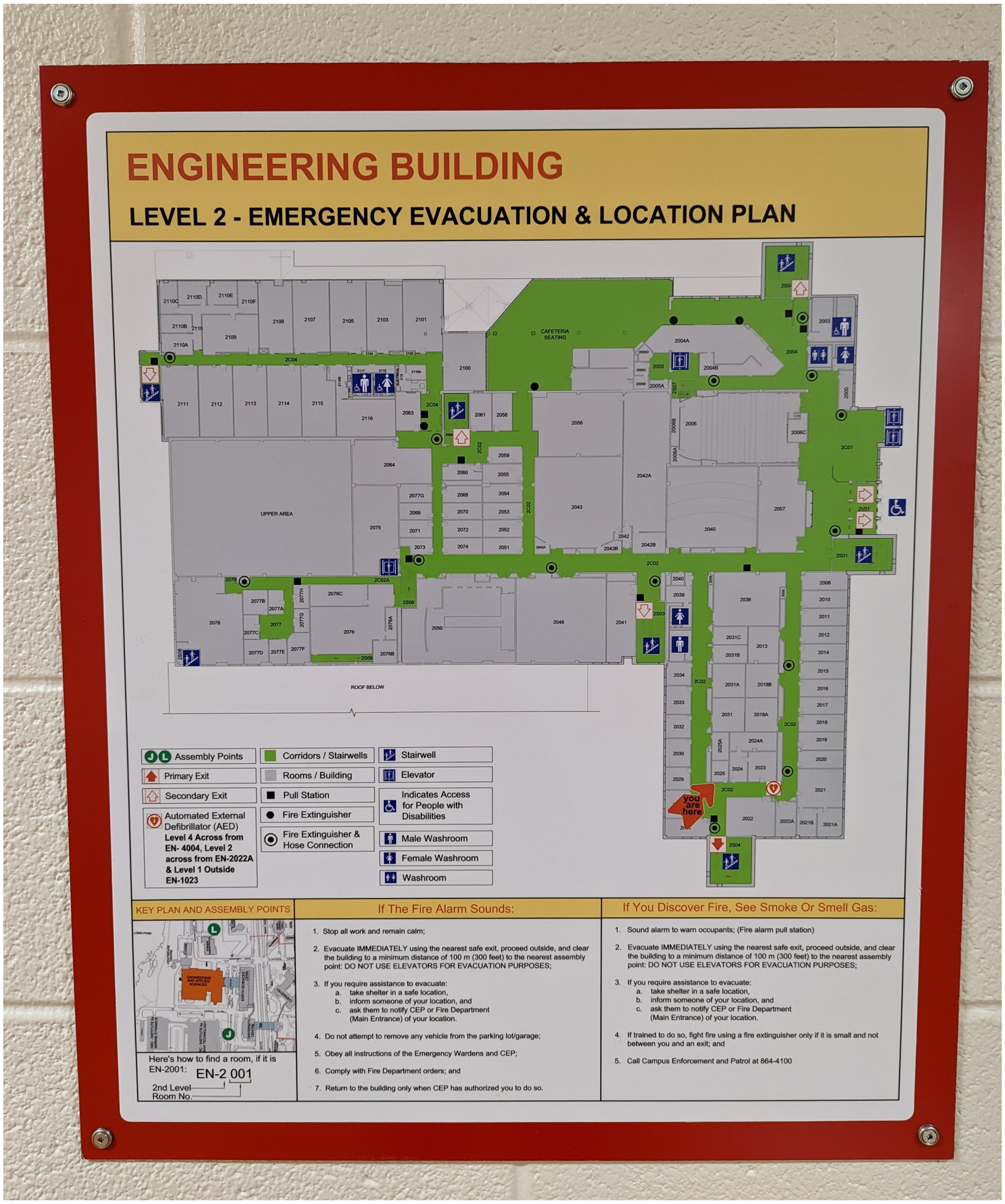

Figure 3.1: S.J. Carew Engineering And Applied Science (EN) Building – Level 2 Floor Plan

3.1.2 Landmarks

The following landmarks of interest will be used in the mapping and navigation process in this project. We will obtain 3D mesh models from some of the landmarks (Level 2 staircase, EN-2013 Visual and Analytic Computing Lab, Level 3 vending machines) in order to demonstrate x-ray visualization during the navigation process. We will select the origins and destinations from these landmarks and create a pathway of spatial anchors between them.

Figure 3.2: Landmarks of interest - (Top Left) level 2 staircase, (Bottom Left) Level 3, (Top Center) Level 2 hallways, (Bottom Center) EN-2013 Visual and Analytic Computing Lab, (Right) Level 3 vending machines

3.2 Tools

3.2.1 Hardware

Microsoft HoloLens 2.

Mapping and navigation of the landmarks are done using the Microsoft HoloLens 2.

3.2.2 Software

3.2.2.1 Unity v2020.3.42f1

Main game engine used to create and build user interface (UI) used in the mapping/navigation process. The game objects can be directly prepared and displayed on the scene. Various C# scripts can be written and attached to game objects (e.g. gesture interaction script). Application can be built and compiled directly in Unity before running on Visual Studio.

3.2.2.2 Microsoft Visual Studio 2022 v17.4.1

Main development environment used to develop various interaction and spatial mapping scripts used in the mapping and navigation process. Also used to deploy application on the Microsoft HoloLens 2.

3.3 Packages and libraries

Mixed Reality Toolkit (MRTK) v2.8.2

The MRTK provides user interface (UI) building blocks and input components like motion interactions and spatial awareness configuration profile The user interface components of the application are developed with various prefabs in the MRTK package (hand menu, buttons, etc).

Azure Spatial Anchor SDK v2.13.3

The Azure Spatial Anchor SDK is used to integrate Azure Spatial Anchors in our application to map the pathway from different locations of interest. The anchors represents precise location of the 3D space and can be saved to disk or to the Azure Cloud, and to be retrieved later during navigation process.

Spatial Awareness System

1. SpatialMapping API
2. IMixedRealitySpatialAwarenessMeshObserver API

Spatial Awareness System on the Microsoft HoloLens provide access to the collection of 3D meshes of the environment during mapping process. We will use the Spatial Awareness library to access the 3D mesh of the environment and configure the mesh observer during the mapping process.

3.4 Procedure

3.4.1 Obtain 3D mesh model of landmark of interest using 3D reconstruction

In order to visualize a landmark (e.g. a lab), we need to obtain a 3D model of the landmark that we can save it in the disk or cloud database before retrieving it and rendering the model when the location (where the landmark is) is selected as the destination during navigation. We propose using video-based 3D reconstruction technique like Structure from Motion (SfM) algorithm to map out the interior layout of the landmark into a 3D mesh model as it is more efficient than image-based reconstruction techniques. Microsoft MRTK provides a library for video-based 3D scene reconstruction using the HoloLens through its Spatial Mapping API. We can generate a 3D mesh model of the landmark by just 'scanning' the room with the HoloLens. In this implementation, we use *medium* for the *Level of detail* settings in our spatial observer configuration profile which defines the triangles / cubic meter used in the generation of the 3D mesh model. The rationale of choosing *medium* instead of *coarse* level of details is that we only want prominent features of a landmark (e.g. outline of the walls, floors, staircases) to be mapped into our mesh model and reduce the level of noise. Additionally, in order to achieve a spatially correct augmentation of the

3D mesh model on the scene, we also need to know the spatial relationship between the environment and the model in the scene. In previous works like [4] which uses marker-based localization to find the homography between the scene and the model in order to overlay the 3D model. Continuous spatial mapping feature in Microsoft HoloLens allow spatial map data to be stored in the device as a point cloud. Over time, the 3D point cloud of the environment will be generated. Instead of generating a 3D model of the room and then finding the homography between the model and the scene, we will take advantage of the point cloud generated by the HoloLens to access and render the selected mesh via IMixedRealitySpatialAwarenessMeshObserver API.

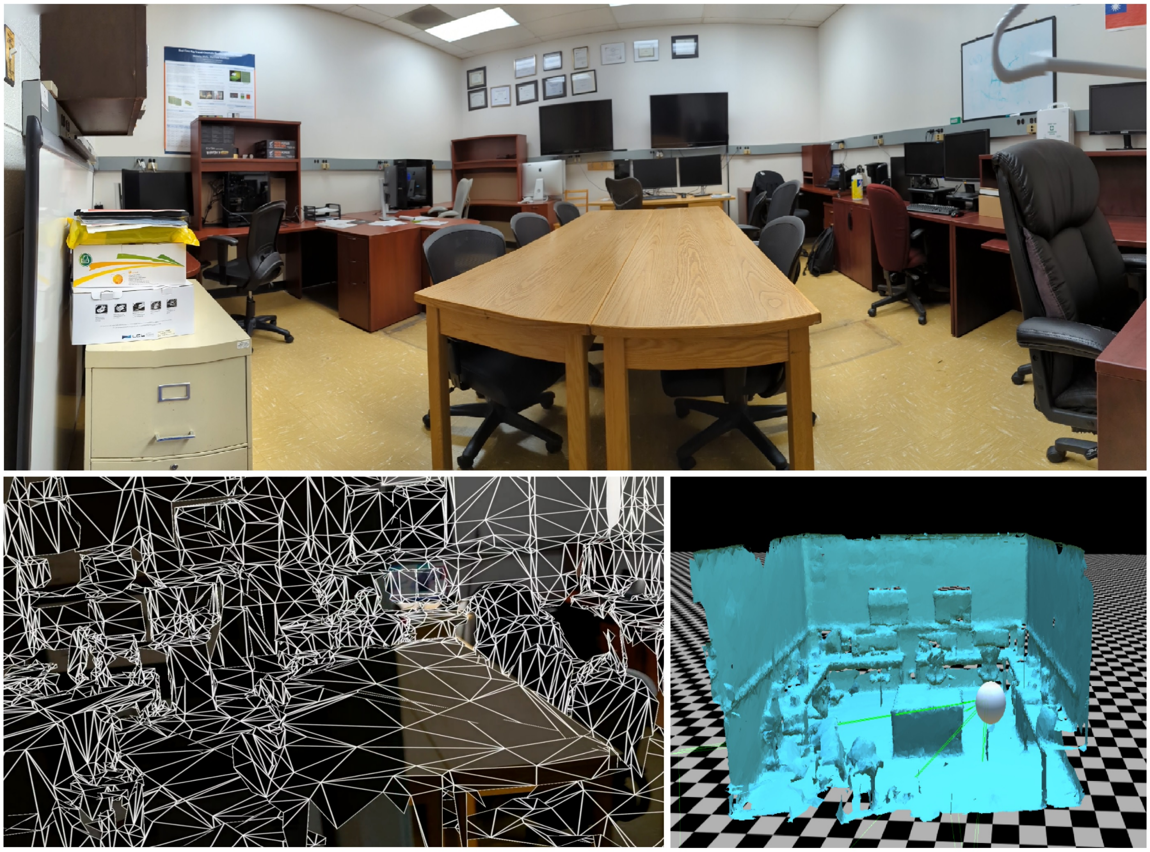

Figure 3.3: Spatial mapping of EN-2013 Visual and Analytic Computing Lab - (Top) Paranoma picture, (Bottom Left) 3D mesh model, (Bottom Right) Point cloud

3.4.2 Path creation and way-finding

After generating 3D mesh models for all the landmarks of interest, we need to map the pathway from various origins to these landmarks of interest. We will use cubes as our game objects to form a pathway for the users to follow during navigation. The rationale for using cubes as opposed to arrows or transition lines is that the users can see the pathway through the walls and gauge the degree of elevation of the terrain (e.g. going from level 2 to 3). We will save a list of cubes which represents the pathway from one location to the specified destination. The shortest path for each destination is decided beforehand and plotted manually during the mapping process as finding the shortest path (taking a list of created anchors and output a list of anchors that represent the shortest path) is beyond the scope of the project. We want to focus on visualizing landmarks and pathway that are occluded by physical objects. Microsoft Azure Services provides Azure Spatial Anchors (ASA), which are precise points of interest in 3D space. Each cube is assigned an anchor ID and will be saved into the disk during the mapping process and retrieved later during navigation. In our implementation, a cube is generated, assigned an anchor ID, and save to the list, and then place on the scene whenever the user 'airtap' in the 3D space.

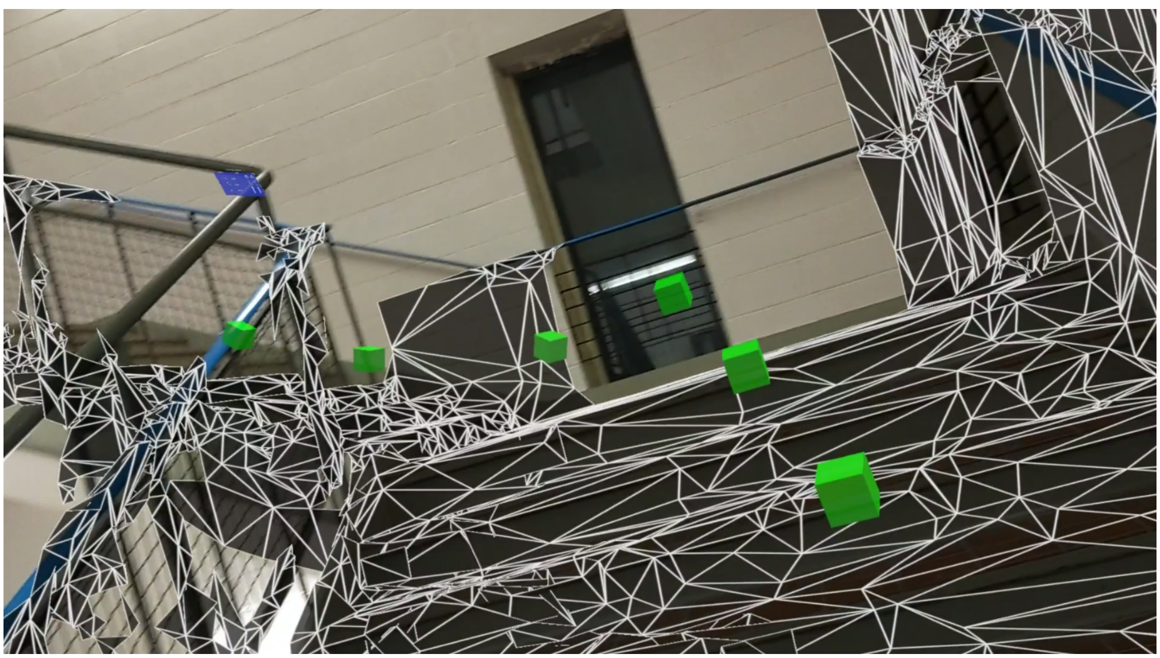

Figure 3.4: Plotting a pathway with spatial anchors

3.4.3 User Interface for navigation and mapping controls

We utilize MRTK examples prefabs to build our user interface (UI) in our application to allow users control the mapping process. The 'Stop mapping' button allows user to stop the spatial mapping process (mapping of landmark into 3D mesh model). The 'Start mapping' button allows the user to resume the spatial mapping process. The hand menu will teleport from another location to directly in front of the user whenever the user lift the palm of their hand, and it can be placed on the 3D space whenever the hand is dropped.

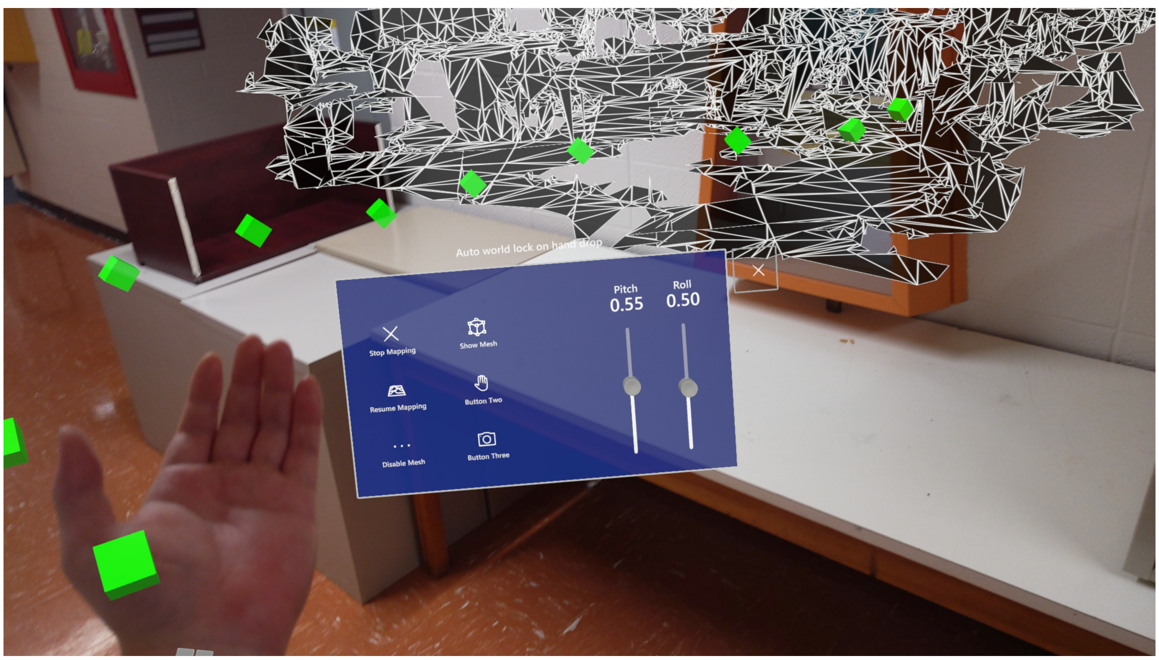

Figure 3.5: Gesture-activated hand menu with mapping controls

Chapter 4

Results and Discussion

Spatial anchors and a 3D mesh model of the destination are generated when the user choose the origin and destination.

4.1 Origin: Level 2 hallway, Destination: EN-2013 Visual and Analytic Computing Lab

Figure 4.1: X-ray visualization of Visual and Analytic Computing Lab with spatial anchors pathway

From Figure 4.1, you can see that the user can intuitively gauge the depth and location of the lab even when the lab is occluded by multiple layers of walls and furniture in the hall. Furthermore, the interior layout of the lab and its furniture can be seen outside the room, allowing the user to gain better spatial understanding of the space without going inside the room. Noise is observed in the 3D mesh model of the lab. One approach is to increase the level of details in the spatial mapping process by increasing the number of triangles / cubic squares in order to increase the fidelity of the model. Gaps in the mesh model are also observed in the model. We can increase the amount of time spent on the scanning process and number of iterations spent on scanning the same surface in the room. However, more work needs to be done to optimize the spatial mapping process.

4.2 Origin: EN-2013 Visual and Analytic Computing Lab, Destination: Level 3

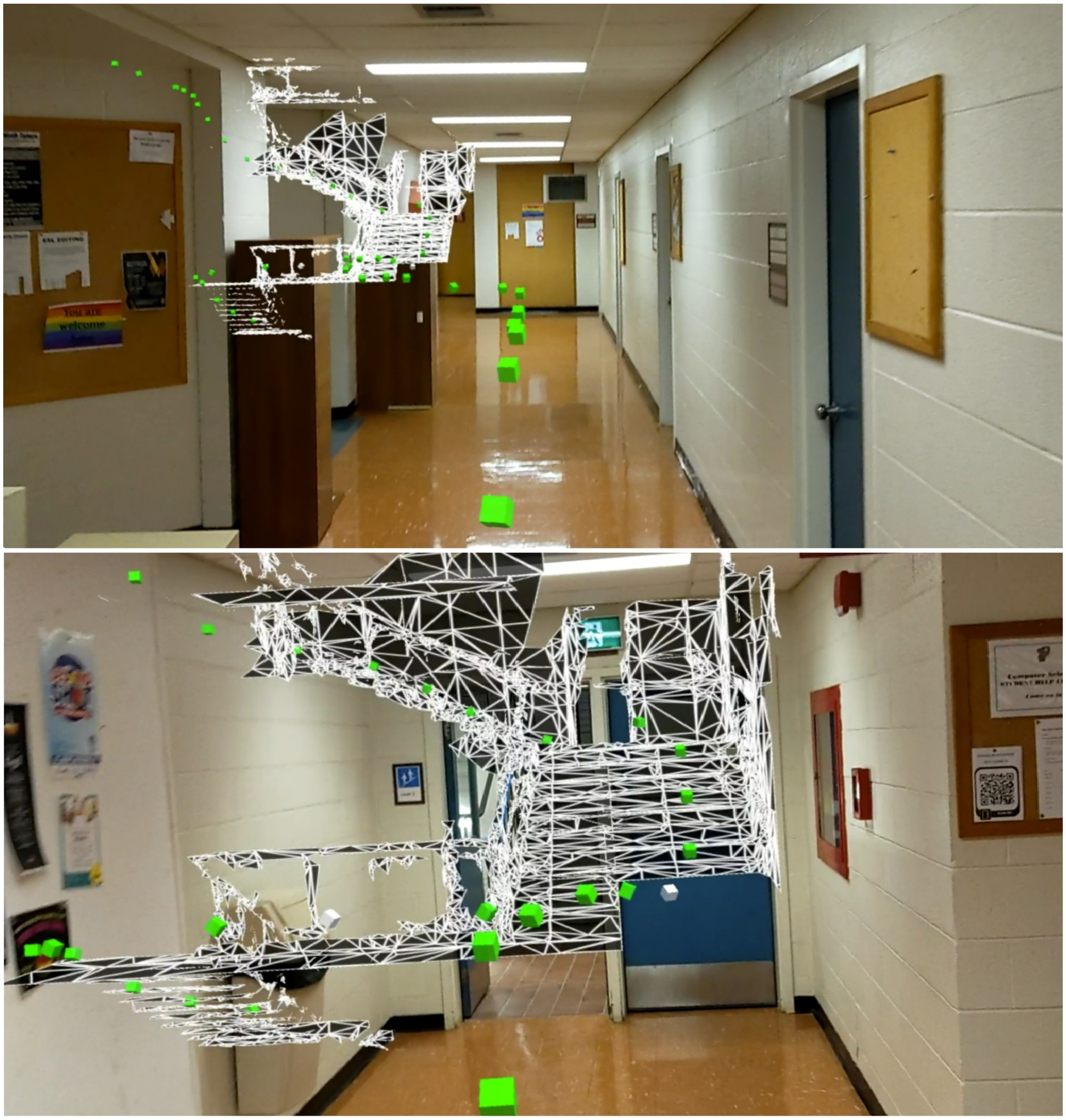

Figure 4.2: X-ray visualization of a 3D mesh model level 2 staircase landing.

From Figure 4.2, we can see that the nearest exit point (level 2 staircase landing) to get to level 3 is being visualized instead of the destination (level 3 of the building). The rationale is that we want to prevent 'drowning' the user with huge amount of visual cues by rendering the level 3 structure of the building.

4.3 Origin: EN-2013 Visual and Analytic Computing Lab, Destination: Vending machines on Level 3

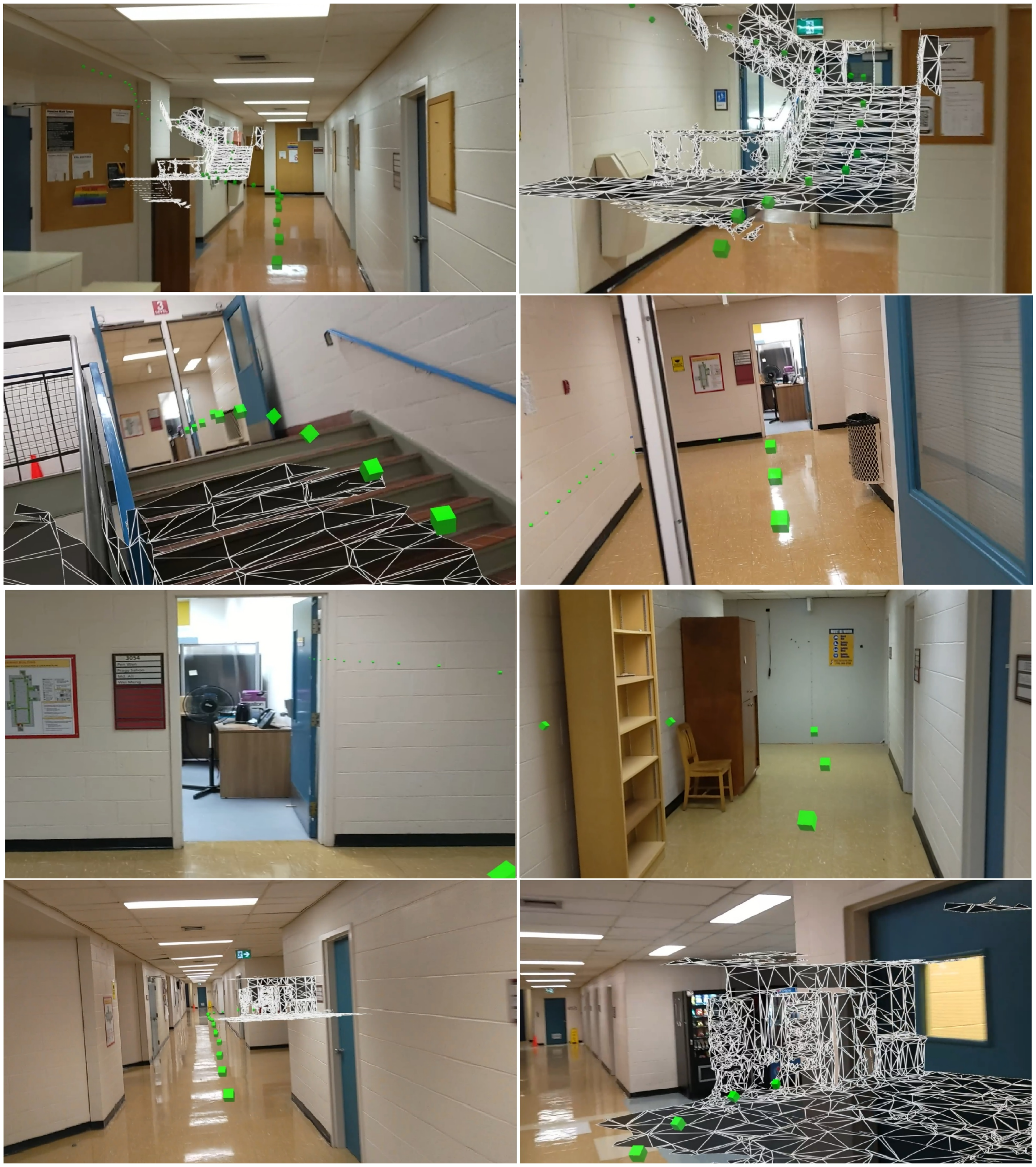

Figure 4.3: Viewing the x-ray visualization of nearest staircase and level 3 vending machines

From Figure 4.3, we can see that the nearest exit point (level 2 staircase landing) to get to vending machine on level 3 is being visualized first before the visualization of the destination (vending machines on level 3). In this example, we want to demonstrate a different workflow in the visualization process when the destination is very far away from the origin. Instead of visualizing the destination from the origin to the destination, we want to visualize the nearest landmark on each checkpoint with the destination being the last checkpoint. The rationale is that if the destination is very far away, the appearance of the 3D mesh model of the destination is too small for the user to decipher the mesh model.

4.4 Origin: Level 3, Destination: EN-2013 Visual and Analytic Computing Lab

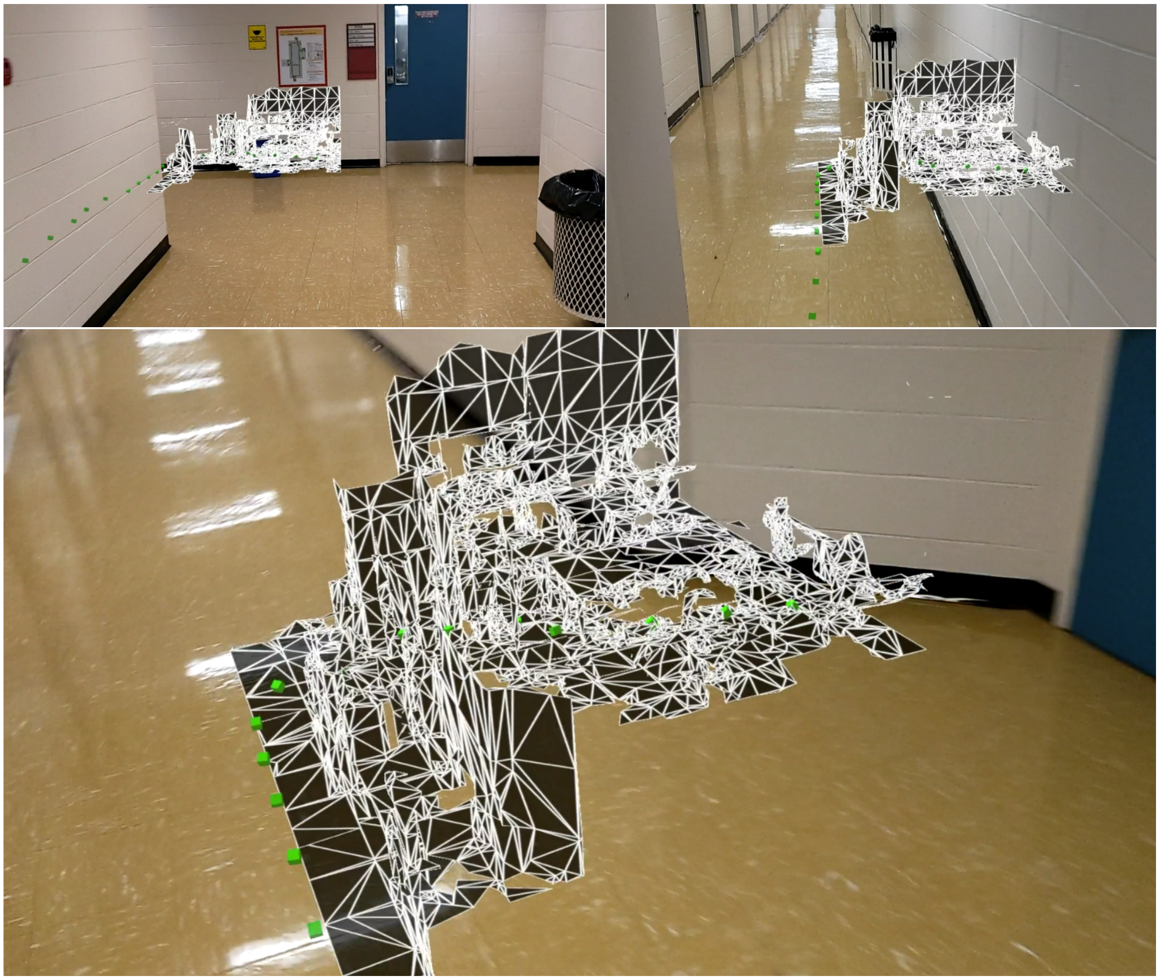

Figure 4.4: Viewing the x-ray visualization of EN-2013 Visual and Analytic Computing Lab from level 3

From Figure 4.4, we can see that the interior structure of the lab is seen from a topographical view which provide visual cue to the user that the destination is one level below the origin.

4.5 Challenges

4.5.1 Environment mapping and scanning

Quality of the 3D mesh models are observed to deteriorate as the mapping data volume gets larger (e.g. when doing 3D scan of the whole building or 3D scan of the pathway to landmark that is faraway from origin). The lighting of the environment and the color of the object/structure play a huge role in the quality of 3D scans. Black objects or low contrast objects are not accurately scanned by the headset. This will affect how the appearance of the 3D mesh model of the landmark, and thus impact the users' ability to perceive or recognize features of the visualized landmark.

4.5.2 Positional shift in 3D model and anchors visualization

Positional shift observed in the rendered 3D mesh model and spatial anchors during navigation and mapping process. Noise in the environment and internet connectivity may have affected the localization ability of the HoloLens and momentarily shift the positions of the 3D objects. The same problem is observed when long distance mapping is done (placing spatial anchors from one origin to a destination is that far away).

Figure 4.5: Positional shift in 3D model visualization - (Left) Hyperloop model at the level 2 lobby of Engineering building. (Top Right) 3D model in the correct position. (Bottom Right) Positional shift in 3D model

Chapter 5

Conclusion and Future Works

5.1 Conclusion

Visualizing 3D mesh model of a landmark provide an effective visual cue to the users to locate the exact position of the landmark that is occluded by multiple layers of objects during navigation. Furthermore, x-ray visualization allows the users to see through walls, thus allowing users to gain better knowledge of the surrounding and thus reducing the amount of time for users to locate the destination.

5.2 Future Works

The identified limitations or encountered problems during this project provide opportunities for future works. The implemented navigation system can be optimized to provide seamless user experience and improved performance or it can be used as a starting point to develop novel applications with extended functionalities in the domain of augmented reality. Some suggestions are formulated below:

1. Implement ghosting into the rendering of 3D mesh models: Although Microsoft HoloLens was not designed as an indoor mapping device, the rendering performance provide a sufficient quality for the users to estimate the depth of the location of the landmark. However, we can improve the blending of the 3D mesh model with the physical environment by introducing ghosting pipeline into the visualization process.

2. Optimizing 3D reconstruction of mesh model: MRTK provide 3 levels (fine, medium, coarse) of details in configurating the mesh observer. 3 levels of details were tested in this project. The mapping process produced either noisy 3D mesh (fine setting) or 3D mesh model with gaps/holes (coarse setting). We need to explore other 3D mesh processing or reconstruction techniques that can increase the fidelity of the generated mesh model.

3. Mapping the whole building: The current implementation only provides navigation between 3 landmarks due to time constraint. We wish to create a 3D mesh model of the entire building and the ability to segment the mesh model of the building into multiple mesh models (mesh model for each lab, staircase landing, exit and entrance points, auditorium, canteen). The big picture is to create a digital twin of the building and only render the specific part of the building based on the destination chosen by the user.

4. UI controls and mapping worfklow: While the hand menu provide a basic UI to control the mapping process, we wish to provide a drop-down menu that would populate the list of destination based on the current location of the user. We also want to change the method we use for mapping the pathway (placement of spatial anchors using airtap) by writing a script to automatically place spatial anchors on the ground by just walking and navigate in the building, with a specifed gap between each anchor.

5. Implement shortest path algorithm for navigation: The shortest path for each landmark is decided before the mapping process in the current implementation. We want to improve the system by integrating shortest path algorithm into the mapping worfklow so that the user can just navigate the whole building and the system will automatically generate a shortest path for each destination from various origins using the generated spatial anchors and save those pathways into the database to be retrieved during navigation.

Bibliography

[1] Toru Ishikawa, Hiromichi Fujiwara, Osamu Imai, and Atsuyuki Okabe. Wayfinding with a gps-based mobile navigation system: A comparison with maps and direct experience. *Journal of Environmental Psychology*, 28(1):74–82, 2008.

[2] Stephen Se, Pezhman Firoozfam, Norman Goldstein, Linda Wu, Melanie Dutkiewicz, Paul Pace, and J. L. Pierre Naud. Automated UAV-based mapping for airborne reconnaissance and video exploitation. In Daniel J. Henry, editor, *Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications VI*, volume 7307, page 73070M. International Society for Optics and Photonics, SPIE, 2009.

[3] Martin Weinmann, Sven Wursthorn, Michael Weinmann, and Patrick Hübner. Efficient 3d mapping and modelling of indoor scenes with the microsoft hololens: A survey - pfg – journal of photogrammetry, remote sensing and geoinformation science, Oct 2021.

[4] P. Hübner, M. Weinmann, and S. Wursthorn. Marker-based localization of the

microsoft hololens in building models, Sep 2018.

[5] Stefanie Zollmann, Tobias Langlotz, Raphael Grasset, Wei Hong Lo, Shohei Mori, and Holger Regenbrecht. Visualization techniques in augmented reality: A taxonomy, methods and patterns. *IEEE Transactions on Visualization and Computer Graphics*, 27(9):3808–3825, 2021.

[6] Stefanie Zollmann, Denis Kalkofen, Erick Mendez, and Gerhard Reitmayr. Image-based ghostings for single layer occlusions in augmented reality. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pages 19–26, 2010.

[7] Mark A. Livingston, J. Edward Swan II, Joseph L. Gabbard, Tobias H. Höllerer, Deborah Hix, Simon J. Julier, Yohan Baillot, and Dennis Brown. Resolving multiple occluded layers in augmented reality. In *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '03, page 56, USA, 2003. IEEE Computer Society.

[8] Woranipit Chidsin, Yanlei Gu, and Igor Goncharenko. Ar-based navigation using rgb-d camera and hybrid map. *Sustainability*, 13(10), 2021.