Université Grenoble Alpes

Master of Science in Informatics at Grenoble

Human Computer Interaction

A Comparative Study of Mobile AR, AirTag and Unaided Search for Indoor Object Finding

---

Authors:

MANGONAUX Leila, SIM Daro

Submitted to:

Prof. Céline Coutrix
Dr. Vincent Lambert

January 15, 2026

I Introduction

Every year, the average American spends 2.5 days looking for lost items and, in the worst case, having to replace them [2]. The question arises, how this search time could be minimized, reducing both the lost time and frustration that come with lengthy searches. Simultaneously, the emergence of computer-vision based augmented reality (AR) technologies has opened opportunities for new interaction methods with computers, including smartphones but remains yet to be systematically explored in the context of lost object finding. Therefore, we ask: **How does visual AR guidance compare to audio-based (AirTag) and unaided search for locating lost objects in an indoor environment?** In this work, an AR mobile application that can assist users in their search for lost items is proposed and evaluated in comparison with existing solutions.

II Related Work

2.1 AR Tracking Methods

Object tracking in AR consists of two phases: During the *registration phase*, the system determines the position and orientation of the device with respect to the real world. Then follows the actual *tracking phase* in which the user's position is updated upon movements [7].

Two primary tracking technology types exist: Marker-based trackers are based on fiducial markers (visual points of reference) or tags (e.g. RFID tags). While these provide high accuracy and stability in tracking at a low cost, they inherently require pre-installation of the markers, limiting the flexibility [7].

Marker-less tracking can be based on sensors or computer vision technologies. Sensor-based solutions require additional hardware and, similarly to markers, may have to be installed in a fixed room before usage. Moreover, each type of sensor has its own advantages and disadvantages. For instance, magnetic sensors allow a full range of motion but limited positioning range [7].

Computer-vision-based solutions rely on optical sensors to capture the surroundings. Objects can be tracked based on previously created models, leading to higher robustness. Alternatively, simultaneous localization and map building (SLAM) has been developed for unknown environments [7].

2.2 Existing Solutions for Object Localization

A review of existing solutions to find lost objects rapidly shows that tag-based solutions are currently the most common. Multiple producers offer tags, namely Apple's AirTags [1] and Samsung's SmartTags [2], that must be attached to objects prone to be lost. They use the Bluetooth Low Energy (BLE) protocol and rely on other devices of a defined operating system to capture the BLE signals and share the GPS coordinates for localization. The tracking accuracy and localization frequency therefore largely depend on the adoption of the corresponding operating system (and users' opt-in to activate location tracking) [4]. Therefore, and due to inaccuracies in GPS localization, the usage of this functionality is often not suited for indoor search. The tags can thus also be prompted to ring for an audio-based search.

Samsung's SmartTag+ additionally offers a mobile AR app to visually indicate the tag's location, developed specifically for use cases in which the object of interest is out of sight [5]. Similarly, Boroushaki et al. [1] developed an AR system using a head-mounted device (HMD) to locate RFID-tagged items.

However, all of these systems require the objects to be tagged in advance. This prevents

[1] `https://www.apple.com/airtag/`
[2] `https://www.samsung.com/uk/mobile-accessories/galaxy-smarttag2-black-ei-t5600bbegeu/`

the search for objects such as glasses that cannot be tagged, as well as spontaneous search for untagged objects.

Mitigating these disadvantages, multiple vision-based solutions have been developed. Yagi et al.'s [9] GO-Finder is a wearable camera that registers hand-held objects such that users can browse the recordings of object interactions via a mobile app. Similarly, Takahiro et al. [8] proposed an HMD system that films pre-registered objects. Users can then view the recordings to identify where they last placed the objects of interest. Oshimi et al. [6] also developed an HMD system which indicates the object location via augmented reality.

While these solutions avoid the discussed weaknesses of tag-based solutions, they require constantly wearing HMDs or cameras to register the interactions with objects. This can be considered cumbersome and unlikely in every-day life. Moreover, the systems cannot capture the location of objects moved off-camera, for example by other people.

In summary, there is currently a lack of vision-based solutions that function both tag-less and without constant registration.

III Prototype

The mobile AR app was built in Unity since this allowed for AR tracking, localization, and navigation. It shows visual AR cues that lead the user from their durrent position to the lost object's location. The project is available on Github [3] and its system implementation is described in further detail in section A.

The user starts by opening the app and holding the phone up toward the room. To let the app capture the space and determine the user and object's position in the room, the user moves the phone around while looking through the camera view. When localization succeeds, the app indicates the location of the lost objects with markers. The user can select an object to search first. The app then starts navigation and draws a clear path on the floor using arrows that point in the direction the user should walk, as shown in figure 10. As the user walks, the arrows update to stay aligned with the user's position and the target. The user can either stop early by tapping “Stop Navigation” as soon as they see the object and know where it is, or continue following the arrows until reaching the destination where the app automatically detects that the target location has been reached and ends the navigation (see figure 11). This process can then be repeated for the remaining targets.

However, due to time constraints, the current prototype does not offer real-time object detection. Therefore, both the room and the object locations must be captured and saved before use.

IV Evaluation Study

A user study was conducted to evaluate the proposed solution and compare it with a state-of-the-art tag-based solution, using AirTags. Only the audio feature of the AirTag was used since this is better suited for indoor environments. Moreover, a control condition was introduced with no helpers. The three conditions were evaluated within subjects, that is, each participant performed the task under all three conditions sequentially.

The users' task was to find three objects (a key, watch and powerbank) in a closed room as fast as possible. Since the prototype did not allow for real-time object detection, a Wizard-of-Oz (WoZ) methodology was selected for which the object detection was mocked by previously setting the object locations in the app. The need to first "scan" the room to detect the objects could also be mocked via the lengthy registration phase needed to localize the user within the room (see section III).

The locations of the objects varied for each condition. To minimize bias, both the order of

[3] `https://github.com/Daro-S/ObjectARIndoorFinder.git`

conditions and the combination of condition and object location rotated between participants following a Latin square, as shown in Table 1.

For each experiment, the search time, number of (wrong) searched locations and, if applicable, the number of times the helper was used, were recorded. For this purpose, participants were asked whether they agreed to being filmed. All except one participant agreed. Thus these metrics were measured in real time for this participant.

After the experiment, participants were also asked to answer a multiple choice questionnaire to examine their satisfaction and perceived usability of the different solutions.

4.1 Participants

A total of nine participants took part in the experiment. All were male students (8) or researchers (1), aged between 20 and 25 (mean 23) years. None of the participants reported uncorrected hearing, sight or mobility impairments. 8 out of 9 participants had used AR before, though none did so regularly.

4.2 Environment and Object Locations

A university study room was selected as study location since it provided sufficient space to distribute the objects such that they could be found in reasonable time without standing out immediately. However, the room was sparsely furnished. To add visual distractions, a set of everyday objects was placed around the room, such as rucksacks, hats, laptops etc. This created options to "hide" objects among other items.

Three sets of object locations were selected such that they were deemed equally difficult: Each location set was composed of two objects placed on the table and a third beside or under the table. Moreover, all objects were placed such that they were (partly) visible from the entrance since occluded objects would not be detectable if the object detection were implemented.

V Results

A comparison of the average search time by condition, as shown in figure 1, reveals noticeable differences with users taking 61.4% longer with the AR app (112.8s) than without helpers (69.9s). In contrast, the AirTags (55.9s) reduced the average search time by 20%.

Similar results can be observed in figure 2 with regards to the accuracy of each condition: The AirTags again lead with 2.9 wrong search locations whereas users searched in 3 to 4 times the amount of false locations without helpers (6.2) and the AR app (8.6).

Further analysis reveals a significant correlation between the number of wrong search locations and search time correlate for the AR app (Pearson correlation coefficient $r = 0.8542$ with $p = 0.0034$) which could not be observed for the other conditions, as can be seen in figure 3. The longer participants searched, the more wrong locations they tended to check when using the app. With AirTags, the number of wrong search locations stayed below 10, and the search time did not significantly correlate with the number of mistakes made. Similarly, some participants in the no-helper condition searched quickly but still checked many wrong places.

Similarly, a significant correlation between the number of times the helper was used and the search time ($r = 0.8235$, $p = 0.0064$)/number of wrong locations ($r = 0.7787$, $p = 0.0134$) was found. The same could not be observed with the AirTag, see figure 5.

Qualitative analysis of the experiment recordings shows that users without available helpers moved around quickly, covering a large part of the room in a short time. In comparison, with a helper, the participants moved slower, regularly pausing to use the helper. Comparing the search time measured per search location shows that participants spent the least time per location without helpers (8.1 seconds) whereas this time increased by approximately three to four times for the AirTag and AR app, respectively (figure 4).

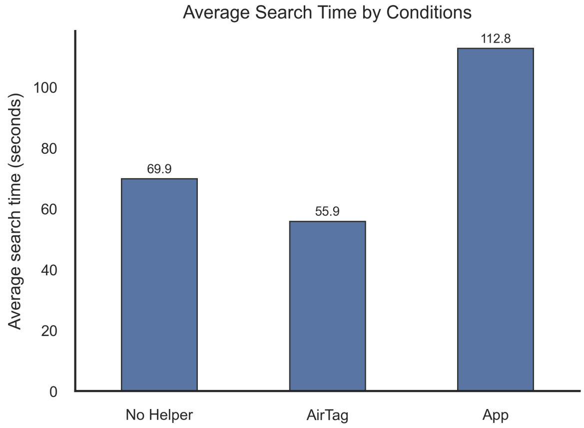

Figure 1: Average search time by conditions

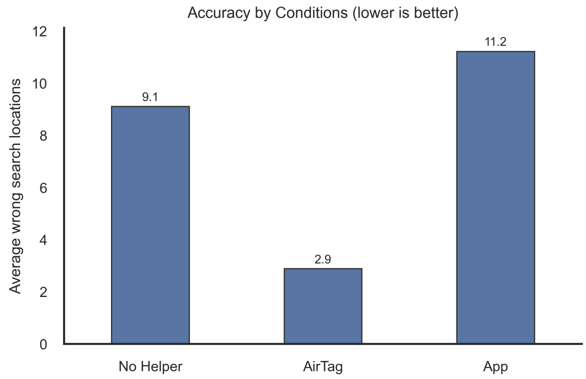

Figure 2: Average wrong search locations by conditions

Furthermore, it could be observed that, when not finding the object in the expected location (which was often expressed verbally), users continued to use the app while searching many locations within a small radius, partially double-checking the same locations.

Finally, it must also be noted that a comparison of object locations revealed noticeable differences. Although the locations were designed to be comparable (see section 4.2), analysis of the corresponding search times by condition indicates significant variance in the difficulty to find objects. Location sets 2 and 3 show similar patterns with the app requiring more than two minutes search time on average while searches without helpers or AirTags averaged at less than half the search time, figure 6. However, location set 1 exhibited a noticeably different distribution in search times. While the average search time with AirTags remained within the same range of around minute, the search without helpers now took longest with a mean of two minutes and the average search time with the app was nearly halved to 68.7 seconds.

5.1 Satisfaction and Workload

The user satisfaction and usability survey revealed no differences in mental demand or perceived complexity between conditions. Both helpers were rated as easy to use by most participants (6 for AirTag, 8 for AR app) and equally helpful (6 each).

All participants agreed that most people would easily learn to use it, compared to only five for the AirTag. However, the AirTag was strongly preferred for accuracy, with seven participants rating it as accurate versus only one for the AR app.

Finally, five participants reported increased stress and frustration when using the AR app, compared to two with AirTags and two without helpers, as shown in figure 7.

VI Discussion

6.1 Discussion of Results

The findings show that providing help during object search can change how people move, decide, and trust guidance, but not all forms of help lead to better performance. The key difference between the helper conditions lies in how much effort they require from users and how clearly they guide the search.

The AirTag performed better because it gave a sound cue with almost no effort. People could keep looking at the room and moving while using the sound to choose where to look next.

In contrast, the AR app introduces additional demands on attention. Users must divide their focus between the physical environment and the screen which can slow down movement and interrupt search flow. When the guidance is not fully accurate, users may rely on it too strongly, repeatedly checking nearby locations instead of expanding their search.

As a baseline, the no helper condition indicates a different limitation: Without guidance, participants move quickly and explore many locations, leading to a high number of mistakes. In this condition, participants favor speed when no support is available, even when this leads to low precision, because they have no clue where to start and must treat many locations as equally likely.

The room layout and object placement also shape how people search and how much a helper can support. When many locations look similar or the space is cluttered, participants cannot quickly decide which places are most likely, so they spend more time checking and re-checking nearby spots. In that situation, any helper cue that is unclear or slightly off can pull participants toward the wrong area and keep them there longer. When locations are more distinct, participants can eliminate unlikely places faster, so even simple guidance supports a smoother and more direct search.

6.2 Limitations

Although carefully designed to minimize the risk of bias or other unrelated effects on the study results, multiple limitations to this work must be mentioned, stemming from the limited scope of this study.

The group of study participants remains small and homogeneous, consisting only of male students and researchers. Potential effects of age, gender or occupation were thus not mitigated and could affect the applicability of the results to a larger population. To expand on this study, we propose recruiting a larger and more diverse group of participants to reduce the risk that the findings reflect behavior specific to a narrow participant group rather than general search behavior.

Furthermore, the lack of object detection functionalities in the app required hardcoding of object locations before the experiments, leading to inaccuracies in the indicated locations. To mitigate this effect in future experiments, we suggest implementing the real-time object detection and using the true object position during each trial and logging the localization error (distance between the real object and the shown AR target) to separate interface effects from errors caused by incorrectly indicated locations.

VII Conclusion and Future Work

This study asked how a mobile AR app compares to audio- and tag-based guidance and unaided search for lost objects. The results show that with regards to efficiency and accuracy, the proposed AR solution could not compete: participants were slower and checked more wrong locations with the AR app than in the other conditions. The AirTag sound performed best because it gave low-effort guidance while participants kept looking at the room and moving, and the no-helper condition was fast but error-prone because participants had no cue for where to start. Overall, low-effort cues that fit natural movement performed better than screen-based AR guidance that demanded attention and depended on accurate localization.

A promising direction for future work is to combine audio-based tags and AR guidance. Audio cues could guide users toward the general area while keeping effort low, and AR could be used only when users are close to the object to provide precise local guidance. This combination could reduce attention load and limit repeated wrong searches. The connection to the audio tag could additionally improve the accuracy of the AR solution and allow for the search of occluded objects.

References

[1] Tara Boroushaki, Maisy Lam, Laura Dodds, Aline Eid, and Fadel Adib. Augmenting augmented reality with Non-Line-of-Sight perception. In *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23)*, pages 1341–1358, Boston, MA, April 2023. USENIX Association.

[2] Pixie Technology Inc. Lost and Found: The Average American Spends 2.5 Days Each Year Looking For Lost Items Collectively Costing U.S. Households $2.7 Billion Annually in Replacement Costs — prnewswire.com. `https://www.prnewswire.com/news-releases/lost-and-found-the-average-american-spends-25-days-each-year-looking-for-lost-items-collectively-costing-us-households-27-billion-annually-in-replacement-costs-300449305.html`, 2017. [Accessed 25-12-2025].

[3] IndoorTracking. The future of indoor navigation: Beyond gps. `https://indoortracking.com/the-future-of-indoor-navigation-beyond-gps/`, 2025. [Accessed 25-12-2025].

[4] Hyunseok Daniel Jang, Hazem Ibrahim, Rohail Asim, Matteo Varvello, and Yasir Zaki. A tale of three location trackers: AirTag, SmartTag, and tile. January 2025.

[5] Samsung Newsroom. Introducing the new galaxy smarttag+: The smart way to find lost items. `https://news.samsung.com/global/introducing-the-new-galaxy-smarttagplus-the-smart-way-to-find-lost-items`, 2021. [Accessed 25-12-2025].

[6] Hiroto Oshimi, Monica Perusquía-Hernández, Naoya Isoyama, Hideaki Uchiyama, and Kiyoshi Kiyokawa. Locatar: An ar object search assistance system for a shared space. In *Proceedings of the Augmented Humans International Conference 2023*, AHs '23, page 66–76, New York, NY, USA, 2023. Association for Computing Machinery.

[7] Toqeer Ali Syed, Muhammad Shoaib Siddiqui, Hurria Binte Abdullah, Salman Jan, Abdallah Namoun, Ali Alzahrani, Adnan Nadeem, and Ahmad B Alkhodre. In-depth review of augmented reality: Tracking technologies, development tools, AR displays, collaborative AR, and security concerns. *Sensors (Basel)*, 23(1):146, December 2022.

[8] Takahiro Ueoka, Tatsuyuki Kawamura, Yasuyuki Kono, and Masatsugu Kidode. I'm here!: A wearable object remembrance support system. In Luca Chittaro, editor, *Human-Computer Interaction with Mobile Devices and Services*, pages 422–427, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.

[9] Takuma Yagi, Takumi Nishiyasu, Kunimasa Kawasaki, Moe Matsuki, and Yoichi Sato. Go-finder: A registration-free wearable system for assisting users in finding lost objects via hand-held object discovery. In *Proceedings of the 26th International Conference on Intelligent User Interfaces*, IUI '21, page 139–149, New York, NY, USA, 2021. Association for Computing Machinery.

Appendix

Table 1: Counterbalancing of condition and location across the three experimental runs (within-subjects). Condition A = AR app, Condition B = AirTag sound, Condition C = no helper.

Participant	Experiment 1	Experiment 2	Experiment 3
1	Condition A, Location 1	Condition B, Location 2	Condition C, Location 3
2	Condition B, Location 2	Condition C, Location 3	Condition A, Location 1
3	Condition C, Location 3	Condition A, Location 1	Condition B, Location 2
4	Condition A, Location 2	Condition B, Location 3	Condition C, Location 1
5	Condition B, Location 3	Condition C, Location 1	Condition A, Location 2
6	Condition C, Location 1	Condition A, Location 2	Condition B, Location 3
7	Condition A, Location 3	Condition B, Location 1	Condition C, Location 2
8	Condition B, Location 1	Condition C, Location 2	Condition A, Location 3
9	Condition C, Location 2	Condition A, Location 3	Condition B, Location 1

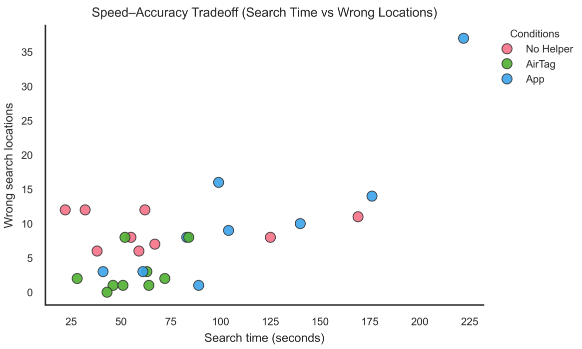

Figure 3: Speed and Accuracy tradeoff by Conditions

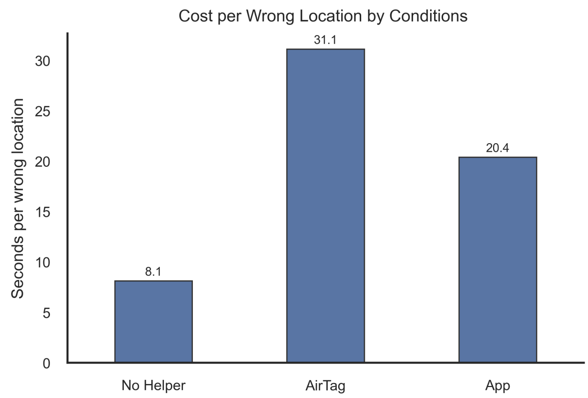

Figure 4: Average search time cost per wrong search location across conditions.

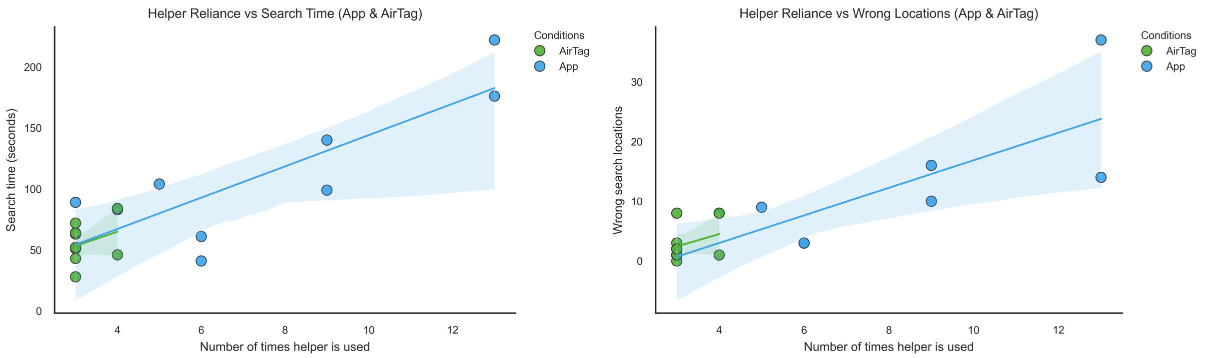

Figure 5: Association between helper reliance and (a) search time and (b) number of wrong search locations.

Figure 6: Average search time by condition for each object location.

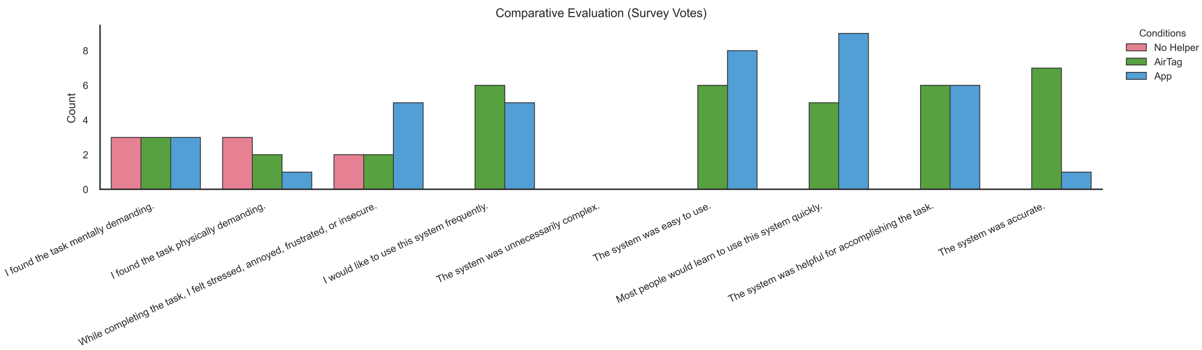

Figure 7: Survey Responses: The questions were multiple choice. The first three questions offered all three response options and the remaining question only offered the AirTag and app.

A Implementation Detail

The overall system workflow, from localization to path visualization and confirmation, is shown in figure 8. A main technical issue with normal mobile AR is that it does not remember an indoor environment. Each time we restart the app, the AR session resets and the phone loses the old map. That means virtual content can drift, and we cannot reuse a previous scan. To solve this, we integrated the Immersal SDK [4]. Before building the app, we prepared and scanned the experiment room and created a sparse point-cloud map, figure 9.

This map stores stable visual features of the space and is stored in the cloud database. When the user opens the app, Immersal's Visual Positioning System (VPS) compares the live camera view to the stored map and returns the phone's position and rotation in that same room. This gives us repeatable indoor localization.

We rely on VPS rather than GPS because GPS is not accurate indoors and cannot locate someone at room scale [3]. VPS uses the camera and the pre-scanned map, so it can place AR content in the correct spot inside a building. Once we know where the user is in the mapped room, we compute a path to the target using Unity's NavMesh system [5]. We baked a NavMesh for the room so the app can tell which areas are walkable and which areas are blocked. We marked walls as obstacles, so the path goes around them instead of through them. The app then draws this path in AR as a sequence of arrows on the floor (Figure Y). As the user moves, we update the path so the arrows keep pointing in the right direction.

Finally, because our study focuses on the guidance experience, we did not implement real-time object detection due to time and resource constraints. Instead, we used a Wizard-of-Oz setup. Before each trial, we manually saved the hidden object's coordinates into the map. During the task, the app guides the user to these preset coordinates and shows a marker when they arrive.

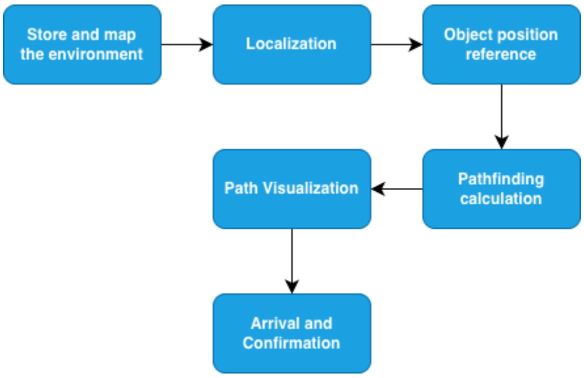

Figure 8: Workflow of the AR-based object search system.

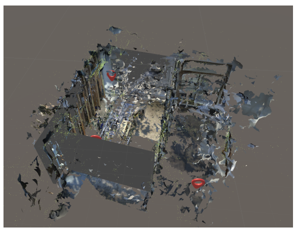

Figure 9: Sparse point-cloud and 3D mesh map of the experiment room.

[4] https://immersal.com/
[5] https://docs.unity3d.com/Packages/com.unity.ai.navigation@2.0/manual/CreateNavMesh.html

1.1 Prototype Details

Figure 10: Path to target.

Figure 11: Arrive at the destination or object found.