Indoor Navigation Systems via Augmented Reality and Reality Capture: From Exocentric to Egocentric Spatial Perspective

Fang Xu, S.M.ASCE[1]; Jason Moats, Ph.D.[2]; Joseph Gabbard, Ph.D.[3]; and Jing Du, Ph.D., M.ASCE[4]

[1]Ph.D. Student, Informatics, Cobots, and Intelligent Construction Lab, Dept. of Civil and Coastal Engineering, Univ. of Florida. Email: xufang@ufl.edu
[2]Director, TEEX Testing and Innovation Center, Texas A&M Univ., College Station. Email: Jason.Moats@teex.tamu.edu
[3]Associate Professor, Dept. of Industrial and System Engineering, Virginia Tech. Email: jgabbard@vt.edu
[4]Associate Professor, Dept. of Civil and Coastal Engineering, Univ. of Florida (corresponding author). Email: eric.du@essie.ufl.edu

ABSTRACT

Conventional spatial information is mainly presented as exocentric perspectives; that is, information is displayed in front of a user, analogous to third person viewpoint. Exocentric perspective presents the location of one object relative to the location of other objects. The recent advancements in mixed reality have expanded the border of spatial information communication by condensing intuitive visualization, localization, and interaction methods into portable devices. The implementation of optical see through (OST) head mounted displays (HMD) and reality capture technologies enable spatial models (e.g., point cloud) to be projected directly at a 1:1 scale on top of the real building, creating an illusion of "seeing through walls." This will give the user an egocentric reference of spatial information, which presents location information relative to the user's own body, and from which the user can plan paths to the desired destination without ingestion and conversion. However, the uncertainties remain in how these applications can improve the efficiency of spatial awareness especially when one has access to both new and conventional spatial perspectives. This paper develops an augmented reality-based navigation system to improve spatial awareness via real-time reality capture and geometric alignment technologies for providing both exocentric and egocentric perspectives. Then a human-subject experiment was performed in a real two-story building at the University of Florida. Corresponding spatial information was presented with pre-scanned point cloud and mesh models using a virtual scene built with Unity game engine and then visualized with a HoloLens 2 HMD in both perspectives. The findings confirmed that the proposed egocentric visualization of building helped achieve better wayfinding performance, and advantages in cognitive load and situational awareness.

INTRODUCTION

Navigation systems have been well developed and adopted for outdoor usage since the introduction of the global positioning system (GPS) in the 1970s. Thanks to the precisely defined GPS coordinates, satellite maps, outdoor navigation can be easily accessible for the public (Panzieri et al. 2002). However, indoor navigation is limited due to the reduced satellite signal strength, the complexity of layouts and occlusion from structures (Simões et al. 2020).

Alternative technologies have been developed to improve indoor navigation using Wi-Fi signals, ultra-wideband (UWB), Bluetooth, and radio frequency identification (RFID) (Zafari et al. 2019). Although these techniques can improve the accuracy of localization and indoor positioning, a problem still presents, i.e., *the design of user interface (UI) that can convey complex spatial information in an intuitive way*. Indoor navigation is quite different from outdoor, especially caused by its spatial complexity when outdoor navigation is usually dealing with 2D layouts and planning routes on the same flattened playground (Ozdenizci et al. 2015). In contrast, indoor structures are made up of multiple levels, stairs, hallways and rooms, which intrinsically inhibit the user from acquiring a “big picture” easily for a healthy spatial cognition. The desired solution is to acquire spatial information with 3D mapping technology in real time and update it to the navigation system as users traverse the indoor area, which indeed provide valuable spatial information to the user, but running a potential risk of information overload if not carefully designed or calibrated. The complexity and occlusion problems of modern buildings add additional difficulty in presenting acquired spatial information as well as landmarks, navigation cues for the user to effectively complete navigation tasks.

Conventional navigation and wayfinding information systems heavily rely on 2D maps, and thus planning for a route is relatively straightforward because of the relative relationship between the user and landmarks can be directly comprehended (Kitchin 1994). Nevertheless, like mentioned earlier, indoor navigation tasks are commonly facing 3D route planning and obstacle avoidance tasks spreading in a 3D space, affecting the spatial cognition of users in a negative way (Hölscher et al. 2006). Even if the target user is well trained for an indoor environment, the high cost of cognition processing would potentially harm the completion of the mission.

To moderate the high cognitive demand caused by a less optimal design of wayfinding information system, as well as merging mapping and displaying technologies for aiding spatial cognition, an augmented reality (AR) (using Microsoft HoloLens 2) based indoor navigation system paired with reality capturing techniques is proposed to present spatial information and navigation cues in an egocentric manner. The remainder of this paper introduces the point of departure, the designs of the system, and a test case to compare the performance between exocentric methods and egocentric methods in wayfinding information presentation.

RELATED WORKS

Navigation is a complex process that involves several cognitive processes. According to the classic Landmark Route Survey (LRS) framework, spatial knowledge is developed with the requirement of understanding the surrounding environment and remembering prominent features before one could plan their route and apply the route knowledge until successfully navigating to the target of interest (Kitchin 1994). This entire process demands cognitive effort if the individual is not aided with any forms of spatial information, which leads to disorientation and loss in practice. Furthermore, for complex indoor structures, landmarks with similar features and repetitive scenes add up to another level of difficulty for one individual to manage their spatial memory. Other studies showed that route knowledge is developed from an egocentric perspective while survey and landmark knowledge are usually learned in an exocentric way (Chittaro and Venkataraman 2006). In addition, survey and landmark knowledge requires spatial memory and an extra cognitive effort to be converted into route knowledge. Due to the limitation of conventional methods of acquiring spatial knowledge and displaying methods, survey knowledge and landmarks couldn’t be presented to the user in an egocentric perspective. In conclusion, spatial perception comes naturally from an egocentric perspective, but traditional ways of providing additional spatial knowledge which are vital for navigation tasks are also introducing extra cognitive load that decreases the efficiency of such navigation systems. With recent advances of spatial understanding and mixed reality technologies, this gap could be mitigated.

AR as a solution of an indoor navigation system has been studied for decades and is gaining attention in recent years because of the development of related technologies as well as the increasing computing power on mobile devices. For example, handheld AR devices are popular in previous studies and the results show that the application of AR could improve localization and wayfinding results (Möller et al. 2014; Mulloni et al. 2011). Since handheld AR devices are displaying virtual objects with solid 2D screens, the margin between the physical world and virtual world is significantly noticeable and could cause disorientation and danger when vision of the user is blocked. In other studies, head mounted displays (HMD) as AR displays are being tested to investigate the impact of AR on indoor navigation (Goldiez et al. 2007; Zhang et al. 2021). These HMDs feature optical or video see-through capability that could project virtual objects on top of the real world and make visualization intuitive rather than handheld devices. Their result shows that mobile AR HMDs could reduce the workload, task difficulty and spatial cognition effort drastically for indoor navigation missions (Rehman and Cao 2016). However, the spatial information provided in these studies is still limited by either the completeness or displaying format. On one hand, even with the help of see-through HMDs, an isolated hologram superimposed on the physical world could be confusing for its unclear relativity with the other objects and a vague estimation of distance. On the other hand, the spatial information provided is based on pre-knowledge and maps are displayed with an exocentric perspective. Previous indoor navigation studies are also limited by their layout of involving only one level of structure, which is seldom the case in practice. A complex multi-level indoor structure in real life scenarios would push the limit of current AR navigation solutions and make the users vulnerable. This study hypothetically considers that an AR navigation system paired with complete 3D structural information and an egocentrically presented spatial information would help users improve their indoor navigation capabilities in comparison to other methods.

SYSTEM DESIGN

Architecture

The proposed navigation system is composed of three major parts, as shown in Fig.1: a reality capturing agent, a computing center, and an AR HMD for displaying virtual objects and interactions. As described in previous studies, spatial information is acquired from a Velodyne Puck 16 channel LiDAR scanner which features fast measuring speed and long range (up to 100 m) (Pu et al. 2021; Xu et al. 2022). The scanning platform is based on a quadrupedal robot agent that can support LiDAR sensor systems and has enough clearance for traversing indoor structures such as stairs (Fig.2). The ranging data acquired from LiDAR is then processed with simultaneous localization and mapping (SLAM) algorithms to form an updating point cloud map for defining a local coordinate system to attach landmarks and localizing the user. The point cloud map is then uploaded to a local computer and rendered by Unity game engine. Then, a HoloLens 2 HMD which accesses the game scene can render the point cloud map for the user in the form of holograms that is superimposed to the physical world. At the same time, the AR

HMD also takes care of the localization of users inside the same physical space, which shares the same coordinate system with the virtual space. The game scene is also interactable for laying virtual objects to markup hazards, blocked entrances, and milestones. Due to the limitation of rendering capacities and computing power of the HoloLens 2 HMD, the point cloud model is rendered on the laptop carried by the user when navigating the physical world. In brief, the combination of virtual and physical worlds is organized under the same coordinate system (game engine) based on the comprehension of the same physical world by sensors from both ends: the user wearing HMD and the ranging sensor carried by robot agent. As a result, a complete closed loop of measuring, mapping, visualizing, and interacting with spatial information is achieved with this proposed system design.

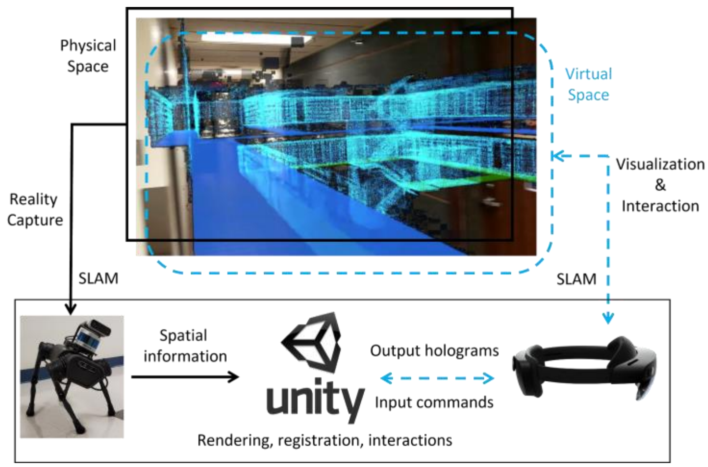

Figure 1. System design

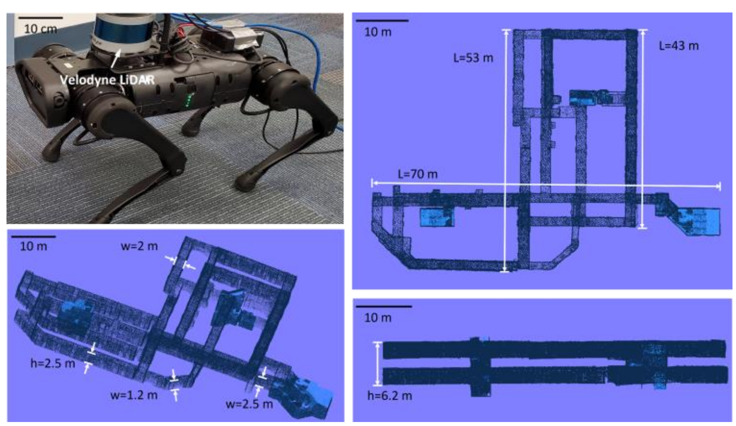

Figure 2. The scanning platform and the case study site

Egocentric and Exocentric User Interface

As shown in Fig.3, the proposed navigation system is designed with two types of user perspective to display spatial information from both egocentric and exocentric perspectives. By using optical see-through AR HMD, both physical and virtual worlds can be visualized at the same time. After the 3D model of indoor structure is precisely measured with LiDAR scanner, and the resultant point could map being rendered with Unity, a direct representation of spatial information can be displayed with 3D model views. The first method, or the conventional exocentric perspective, is a miniaturized 3D model containing the same amount of spatial information as well as the position and orientation of the user (Fig.3b). Due to the potential occlusion caused by multi-level structure and viewing angle, the 3D mini map is also interactable through hand gestures for rotating, panning, and zooming. By using the 3D mini map, the user can have a top-down view of the environment with complete spatial geometry information as well as monitoring tracked objects and placed markers.

The second method, or the egocentric perspective, includes two UI methods. The first of it relies on signages anchored to the environment, such as arrows showing the routes (Fig.3c). The second method of it is named "Global View", which features a 1:1 scaled holograms of the scanned point cloud model as well as desired virtual objects superimposed to the physical world (Fig.3d). With the 1:1 scaling of geometry and sizing of virtual objects, another advantage of this method is a natural estimation of distance and relativity between objects. Global View ensures an intuitive presentation of spatial information and navigation cues so that an egocentric situational awareness can be established.

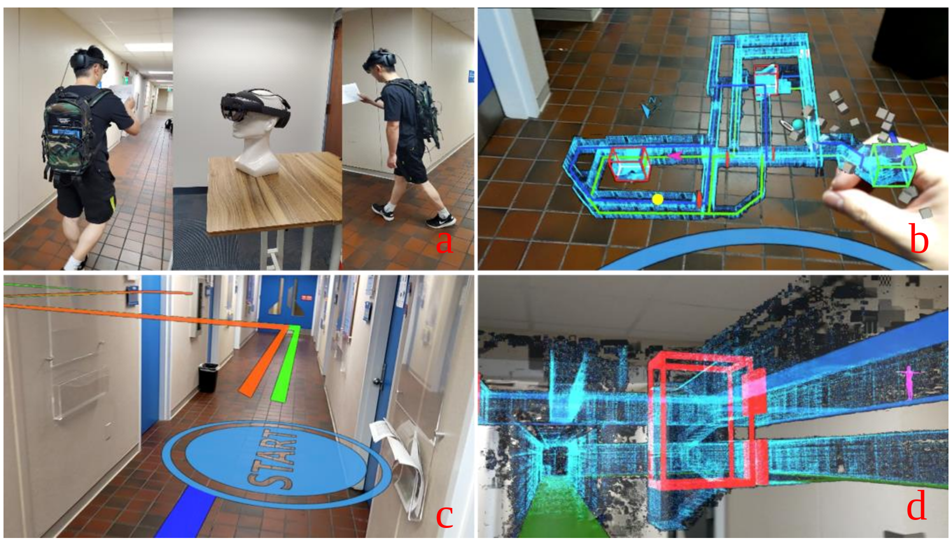

Figure 3. Egocentric and exocentric UI in wayfinding. a) system setup; b) exocentric view – mini map; c) egocentric view – anchored signage; d) egocentric view – FPV global view

Multiuser and Cooperation

In addition, the system is capable of supporting multi-user scenarios and cooperation requirements. By using Photon Unity Networking (PUN), the transformation of designated virtual objects can be synchronized over the internet in real time. PUN is a Unity package for multiplayer games featuring real time cloud service, cross-platform flexibility, and high scalability, which is ideal for the navigation system's requirements. To be noted, the Azure spatial anchoring (ASA) function and world locking tools (WLT) function provided by Microsoft requires it to be built and deployed to the HoloLens to be functional. But the HoloLens 2 is tested to be not capable of rendering complex geometries, especially the scanned bulky point cloud model. PUN's cross platform capability enables the data processing and rendering to take place on the same Unity project with Windows OS, and then projected to HoloLens 2 for visualization.

Within the PUN, connected users are registered under the same virtual coordinate system so that designated virtual objects are assigned with unique identification numbers and then their transform values are synchronized based on the same coordinate system. Spawned objects are assigned with the same ID, and the ownership of the objects is handed over to the user who initiated interaction with the object. In this navigation system, users connected to the same server can visualize each other's location and orientation from any distance at the same time. Placed landmarks, markings for hazards, milestones can also be visualized and interacted with by any desired user. Shared spatial information between users not only guarantees communication but also increases the efficiency of surveying an unfamiliar indoor space.

CASE STUDY

A case study was carried out to validate the hypothetical benefits of the proposed system as well as finding the difference between exocentric and egocentric perspectives for presenting navigation cues (Fig.4).

Figure 4. Case study. a) third person view of a subject navigating the area; b) first person view of a subject at the starting point of a trial

The test subjects (n=31) were instructed to use the proposed navigation system to find a human shaped target (virtual object) inside a 2-level indoor structure on UF campus. There were four conditions tested, including 1) Control Condition (with only a 2D paper map); 2) miniature Map (with a 3D map in hand); 3) Route Condition (with 2D map and routes visualized on the floor); and 4) Global View (with OST at 1:1 scale). Among all, conditions 1 and 2 are considered as exocentric perspectives, and conditions 3 and 4 are considered as egocentric perspectives. The subjects were told to find the target within a 10-min limited time using four types of cues and their performance were recorded and tested with questionnaires after each trial. To replicate the conditions of an emergency situation, virtual obstacles were strategically placed at choke points. In each of the four conditions, subjects were provided with complete spatial knowledge only in different forms or combined forms. For the route condition (condition 3), the subjects were also provided a 2D map with every obstacle marked in place and the route were arrows and lines planted on the floor which represented the shortest path to target before the obstacles were placed.

According to the survey, 35% of the subjects reported that the exocentric cues are more helpful for localizing themselves in the building while 62.5% of the subjects preferred the egocentric cues. As for determining which method was most helpful for navigating towards the target object, 90% of the participants claimed that the egocentric cues were the best. It suggests that egocentric perspective indeed facilitates a better spatial understanding and navigational decision. In addition, we found that the AR as a UI platform played a positive role as well. When provided with the complete AR spatial cues, including condition 2 - miniature 3D map in hand, and condition 4 - global view based on the OST function of AR, all 31 subjects could finish the given task on time (<10 minutes). In contrast, other conditions did not help as much as AR. For example, given 2D paper maps as the main wayfinding information source, even routes on the floor were provided as well, there were still six subjects who failed to find the target within the given time limit. The results showed that even with an advanced indoor navigation system, the performance of individuals would vary significantly due to the way that spatial information was provided. With the OST function of AR HMD providing 1:1 scaled egocentric perspective, users could navigate a complex indoor structure with confidence and efficiency compared to the traditional exocentric maps. Less calibrated UI designs would have a negative impact on the performance, potentially causing confusion for the user.

CONCLUSIONS AND FUTURE WORK

With respect to our research intentions, this paper presents a close loop indoor navigation system backed up with LiDAR scanners for reality capture, OST AR HMD for visualization, real time cloud service for multi-user collaboration and intuitive UI design that features both exocentric and egocentric cueing perspectives. The results indicate that both exocentric and egocentric cues with complete spatial information can benefit indoor navigation tasks. While the egocentric cues can increase the wayfinding and path planning efficiency, and the exocentric cues can help the user with better self-localization. Some studies also indicate that the optimal solution of AR navigation cues would feature the combination of both exocentric and egocentric perspectives, which comes naturally for users who desire a complete level of situational awareness especially in vital scenarios (Melzer and Madison 2020; Shikauchi et al. 2021). The future research plan includes optimizing UI designs to improve user experience with minimal confusions and testing collaboration strategies to achieve better efficiency when handling complex navigation tasks. By integrating advanced sensing, robotics, visualization technologies and intuitive user interfaces, the future engineers and first responders can navigate hazardous and unfamiliar environments in a safer and efficient way. As the prevalence of mixed reality increases, the proposed navigation system holds promising potential to offer a comprehensive solution not only for specialized applications but also for generic scenarios.

REFERENCES

Chittaro, L., and Venkataraman, S. “Navigation aids for multi-floor virtual buildings: A comparative evaluation of two approaches.” *Proc., Proceedings of the ACM symposium on Virtual Reality Software and Technology*, 227-235.

Goldiez, B. F., Ahmad, A. M., and Hancock, P. A. (2007). “Effects of augmented reality display settings on human wayfinding performance.” *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 37(5), 839-845.

Hölscher, C., Meilinger, T., Vrachliotis, G., Brösamle, M., and Knauff, M. (2006). "Up the down staircase: Wayfinding strategies in multi-level buildings." *Journal of Environmental Psychology*, 26(4), 284-299.

Kitchin, R. M. (1994). "Cognitive maps: What are they and why study them?" *Journal of environmental psychology*, 14(1), 1-19.

Melzer, J. E., and Madison, R. W. "Birds do it, Bees do it: A bio-inspired look at wayfinding and navigation tools for augmented reality." *Proc., Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR)*, SPIE, 172-182.

Möller, A., Kranz, M., Diewald, S., Roalter, L., Huitl, R., Stockinger, T., Koelle, M., and Lindemann, P. A. "Experimental evaluation of user interfaces for visual indoor navigation." *Proc., Proceedings of the SIGCHI conference on human factors in computing systems*, 3607-3616.

Mulloni, A., Seichter, H., and Schmalstieg, D. "Handheld augmented reality indoor navigation with activity-based instructions." *Proc., Proceedings of the 13th international conference on human computer interaction with mobile devices and services*, 211-220.

Ozdenizci, B., Coskun, V., and Ok, K. (2015). "NFC internal: An indoor navigation system." *Sensors*, 15(4), 7571-7595.

Panzieri, S., Pascucci, F., and Ulivi, G. (2002). "An outdoor navigation system using GPS and inertial platform." *IEEE/ASME transactions on Mechatronics*, 7(2), 134-142.

Pu, G., Wei, P., Aribe, A., Boultinghouse, J., Dinh, N., Xu, F., and Du, J. "Seeing through walls: real-time digital twin modeling of indoor spaces." *Proc., 2021 Winter Simulation Conference (WSC)*, IEEE, 1-10.

Rehman, U., and Cao, S. (2016). "Augmented-reality-based indoor navigation: A comparative analysis of handheld devices versus google glass." *IEEE Transactions on Human-Machine Systems*, 47(1), 140-151.

Shikauchi, Y., Miyakoshi, M., Makeig, S., and Iversen, J. R. (2021). "Bayesian models of human navigation behaviour in an augmented reality audiomaze." *European Journal of Neuroscience*, 54(12), 8308-8317.

Simões, W. C., Machado, G. S., Sales, A. M., de Lucena, M. M., Jazdi, N., and de Lucena, V. F., Jr. (2020). "A review of technologies and techniques for indoor navigation systems for the visually impaired." *Sensors*, 20(14), 3935.

Xu, F., Xia, P., You, H., and Du, J. (2022). "Robotic Cross-Platform Sensor Fusion and Augmented Visualization for Large Indoor Space Reality Capture." *Journal of Computing in Civil Engineering*, 36(6), 04022036.

Zafari, F., Gkelias, A., and Leung, K. K. (2019). "A survey of indoor localization systems and technologies." *IEEE Communications Surveys & Tutorials*, 21(3), 2568-2599.

Zhang, J., Xia, X., Liu, R., and Li, N. (2021). "Enhancing human indoor cognitive map development and wayfinding performance with immersive augmented reality-based navigation systems." *Advanced Engineering Informatics*, 50, 101432.