Towards Cross-Platform Immersive Visualization for Indoor Navigation and Collaboration with Augmented Reality

Akshay Ayyanchira[1], Elias Mahfoud[1], Weichao Wang[2] and Aidong Lu[1*]

[1]Department of Computer Science, University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, 28173, NC, US.

[2]Department of Software and Information Systems, University of North Carolina at Charlotte, 9201 University City Blvd, Charlotte, 28173, NC, US.

*Corresponding author(s). E-mail(s): aidong.lu@uncc.edu; Contributing authors: aayyanch@uncc.edu; emahfoud@uncc.edu; weichaowang@uncc.edu;

Abstract

Navigation and situation awareness in 3D environments are often required by emergency responses such as firefighters, police, and military soldiers. This paper investigates a collaborative, cross-platform immersive system to improve the team navigation through effective real-time communication. We explore a set of essential building navigation, visualization, and interaction methods to support joint tasks in the physical building environments by leveraging on device sensors and image markers. Our platform also supports efficient exchange of useful visual information to improve the coordination and situation awareness for multiple users performing real-time operations in smart buildings. We have performed a user study to evaluate different devices used in coordination tasks. Our results demonstrate the effects of immersive visualization for improving 3D navigation and coordination for on-site collaboration in real physical environment.

Keywords: Immersive visualization, 3D Navigation, Collaboration, Augmented Reality

1 INTRODUCTION

The techniques of augmented reality (AR) [1–4] have grown continuously since the work of Ivan Sutherland [5]. During the past, the technology was often used in research or prototype systems instead of practical applications. Only until recent years, wearable technology has started to achieve tremendous advancements in many aspects [6], including processing power, display techniques, smaller and more efficient sensors, and more effective software algorithms [6].

Such advancements have started to support immersive systems in a variety of domains, including gaming, education, architecture, facilities management, emergency response and military operations, and many others [7, 8]. Similar to AR, the concept of mixed reality (MR) emphasizes the usages of blending the physical world with the digital world with advanced technologies. Integrating MR with the major advances from other fields such as smart cities and internet of things (IoT), we can create new techniques from these interdisciplinary research fields and revolutionize how we work and collaborate in many professions [9].

For emergency responders, such as firefighters and police, we can collect all available information through smart devices and sensors to improve the understanding of building situations in real-time. We can further integrate machine learning and data analysis algorithms to assist commanders to derive important decisions for operations in the field, such as if the crews should enter a room with potential hazards. For crews working in the fields, we can enable them to share information with others for more efficient collaboration. Such techniques may significantly improve the job safety and performances with efficient data sharing and coordination solutions.

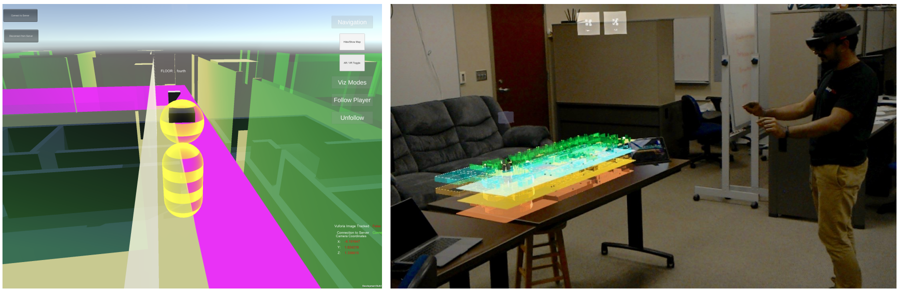

Fig. 1 Immersive visualization and navigation with mobile devices (iPad) and AR HMDs (HoloLens)

This work extends our previous work [10], which presented a cross-platform approach with MR to support the collaboration of multiple users and improve their situation awareness in smart buildings. Our MR platform utilizes mixed-reality devices including AR HMD HoloLens and iOS devices to assist indoor operations such as building navigation and search (Figure 1). We provide a flexible server-client architecture to synchronize user locations and vital information in real-time with two-way communications. We have also developed a set of building navigation, visualization, interaction and collaboration functions for users to access real-time information and perform investigation on-site effectively. Our examples and user study demonstrate that the prototype system supports the team navigation and collaboration tasks and enhances the situational awareness of the physical environment. The results also suggest that such immersive functions are promising for challenging situations including operations in buildings with complex structures or time-critical navigation tasks.

The remaining paper is organized as follows. We first review the related work in Section 2 and present our cross-platform system architecture in Section 3. Section 4 describes our immersive visualization and navigation system and the interaction functions that enable users to perform several common navigation tasks. Section 5 describes the information sharing methods for team collaborations between command center and users in the fields. Section 6 summarizes our user study and Section 7 concludes the paper with future work.

2 Related Work

This section presents the related work from the aspects of augmented reality, collaboration, and cross-device techniques for smart buildings.

2.1 Augmented Reality for Smart Buildings

While augmented reality has been proposed for years, its applications on smart buildings are still not common. The pioneer work from Feiner et al. [11] presented an AR system for showing information about the university campus by overlaid them onto the real environment through a head-mounted see-through display. The system combined sensors of GPS for position, accelerometer for head pitch and roll, and a magnetometer for head yaw. Similarly, Tache et al. [12] presented a AR training tool, which shared operative locations in the field to the command center and displayed that information on a calibrated CAD model. Recently, Huang et al. [13] designed the ARGIS system with a precise registration method, which required users to guide the calibration process and compensate measurement errors through a hand-held device. All of these systems are designed to work outdoors and rely on GPS to provide location information.

The example work for indoor environments also appeared recently. For example , Irizarry et al. [14] presented an AR system to show infrastructure and information by overlaying the Building Information Model (BIM) with the camera view of an iPad. While most of these approaches allowed for interactive data authoring from the server side, but they did not provide operators in the field to communicate effectively with the command center.

2.2 Navigation Assistance

While AR solutions for complex indoor environments are still not fully ready yet due to the requirement of a good 3D model of the environment and the metrics of devices [15], previous methods on navigation assistance have shown that AR navigation system can provide augmented directions from the user's current position to his/her desired location [16]. A systematic review of the literature summarized work on navigation assistance to people with dementia during indoor, outdoor and virtual scenarios [17].

The occlusion has remained as the top challenging problem of navigation assistance [18]. Researchers have studied methods to visualize off-screen points of interest with 3D halo [19], locations with wedge [20], and objects with halo [21, 22]. An interesting direction is to explore multiperspective visualization by bringing simultaneously into view multiple regions of interest [23–25]. In addition, ParaFrustum was proposed to establish constraints on a range of acceptable locations and angles for the user's eyes [26].

Interaction with the environment can be challenging due to the differences between the real and virtual worlds. A study with an immersive gesture interface for 3D map navigation in HMD-based virtual environments had shown improvement on the level of immersion [27]. The Worlds-in-Wedges supported comparative immersive visualization by dividing the virtual space surrounding the user into volumetric wedges [28]. Interactive slice World-in-Miniature combined an interactive multitouch table and a stereoscopic display wall to support navigating and interrogating volumetric data sets [29]. Omnidirectional 3D cursors for mobile augmented reality platforms utilized dynamic perceptual affordances to draw user attention to the target location [30]. Sound support has been integrated and shown to be effective [31].

Evaluation results of navigation assistants have shown that AR-style interaction provided good spatial understanding overall [32–34]. AR supported navigation and flight planning of micro aerial vehicles were shown to be effective on augmenting the work of flight supervision [35]. Combined techniques were shown to be superior in comparison to the unimodal attention guiding techniques and subjective preference by the participants [36]. Evaluation of the ARCore indoor localization technology was also performed [37]. On the other hand, the real environment still received the best performance measures compared to a remote environment via a telepresence system and a virtual simulation of the real environment [38]. Tracking uncertainties in AR navigation was also shown to be challenging due to the lack of suitable visualization methods [39]. The effects of the field of view in AR display on search performance in divided attention tasks were shown to be important [40].

2.3 Collaboration in AR/VR Environments

Both AR and VR have been used to create collaboration systems [1, 2, 41]. The latest surveys of AR have also listed collaboration as an important research direction [41, 42]. An early example from [43] used a stereoscopic display to project onto a see-through display, allowing multiple users to simultaneously view the same spatially-aligned model. Benko et al. [44] developed a system for exploring scaled and full-size representations of an archaeological dig by multiple users. A co-located AR system by Nilsson et al. [45] was to support police and military personnel to perform joint planning tasks with shared organization-specific views. Dong and Kamat [46] introduced a co-located collaboration system with tabletop fiducial AR. Butscher et al. presented an approach of AR above the tabletop for collaborative analysis of multidimensional data [47].

2.4 Cross-Device Techniques

Cross-device techniques have been developed by integrating different devices and framework [48]. The GraSp approach [49] combined spatially-aware mobile displays and a large display wall to support graph visualization and interaction. Horak et al [50] presented the combination of smartwatches and a large interactive display to support visual data analysis. Butscher et al. visualized multidimensional data with a touch-sensitive tabletop and AR headsets [47]. Speicher et al [51] developed a taxonomy of AR system components and identified key challenges and opportunities in making them work together. A challenge of cross-device techniques is to handle the networking and communication among multiple devices. VisHive [52] constructed web-based visualization applications that connect multiple devices transparently. Similarly, our approach contains a client-server networking component to stream data between server and connected clients.

3 Collaborative and Cross-Platform Immersive System

We present example usage scenarios and the architecture of our cross-platform immersive system in this Section, and details of our system and information sharing methods in Sections 4 and 5.

3.1 Usage Scenarios

From our discussions with the emergency responders, the localization of people on the site is one of the most attractive features of AR systems. It can be used to find ways to a targeted location or an exit, which are essential to save both civilians and responders. Previous cases have shown that poor decision making from limited information on site and ineffective communication have caused high risks and resulted in huge casualties [53, 54]. While special localization devices are becoming available, they are often not provided to every emergency responder for various reasons including cost, utility, and the number of devices that are already required to carry. Therefore, more fundamental localization through commercial devices and improvement of communication are really needed.

Such AR systems can be also used in scenarios of civilians when they are at new or complex environments. It is common that people seek help to figure out their locations or paths. One example usage scenario motivates our work is the navigation in our academic building. It is a typical 4-floored office building, however navigation is challenging due to the way that the rooms are numbered non-intuitively. Teachers and students often wonder around the building to find their ways or decide meeting locations with others. Similarly, navigation assistance can be helpful in foreign or crowded environments. Since almost everyone has an AR device such as a phone or even HoloLens nowadays, they can seek for the needed location information conveniently.

As a core component of such AR system, we target at the common scenarios of indoor navigation described above and only require the example AR devices without military-level GPS. Our system is designed to support the following tasks:

- Find locations of users in an indoor environment, which has been modeled and stored in the system.
- Find locations of teammates in real-time.
- Communicate brief information among the team, such as an important location to meet or stay away.
- Mark or select navigation paths for oneself or others to follow.
- Allow different AR devices.
- Use the latest hand/voice interactions for AR systems.

3.2 System Architecture

Our system supports multiple users to perform navigation and collaboration tasks with different types of devices, including HoloLens and iPhone/iPad, as illustrated in Figure 2, Our system consists of a server component, which gathers information about the connected clients and distributes the information to them. Our system architecture supports the center and clients to exchange various information, including the real-time locations of the clients and building status from sensors. We allow users at the center to overview comprehensive information of building, while clients in the field to use different interfaces to synchronize their locations in a building, perform on-site operations, and share information with others. The following describes four important components of our system respectively.

Support for Multiple Devices. We have built our prototype system using the Unity3D engine, and currently support the two types of devices – HoloLens and iOS devices. Specifically, the HoloToolkit.Sharing library allows applications to span multiple devices, and enables holographic collaboration. HoloToolkit.Sharing enables multiple devices to communicate and stay in sync seamlessly in real time. Users can also collaborate with other users who may be in the same room or working remotely. For iOS devices, we use the ARKit 2 integrate shared experiences, object detection and image tracking functions.

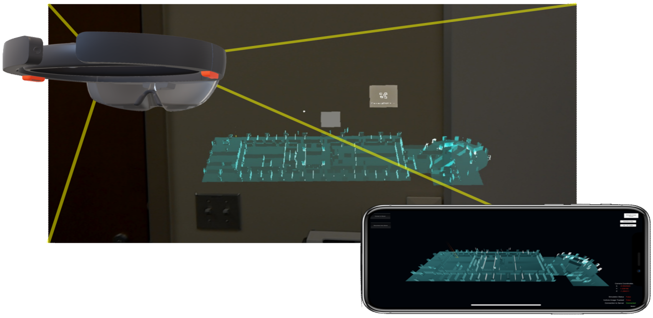

Fig. 2 Application Design for Hololens and iOS devices in our immersive navigation system.

Networking. The networking component of our framework uses UNET, Unity's high-level networking API (HLAPI). We have adopted server/clients mode to keep the components isolated and modular. Our project uses two custom classes inherited from the MessageBase class of Unet, one for client device status sent from client to server and the other for server to broadcast information to connected clients. Message classes can contain members that are basic types, structs, and arrays, including most of the common Unity Engine types.

Coordination of Locations. The problem for collaboration among multiple devices is that each device may come with an independent coordinate system. We cannot simply synchronize user locations directly and assume that all clients share the same coordinate system.

While each device supports a spatial anchor created by the first connected client in the shared system, it is not obvious how to associate the spatial anchor with a 3D CAD model. Instead, we create the common coordinate system through an external, predefined location that has a corresponding point in the CAD model as the common coordinate system. We locate this external predefined point with an image marker from Vuforia, which uses computer vision to locate predefined images in 3D space. Since this library offers strong integration with the Unity engine, we successfully apply it to all our mobile/AR devices. Whenever a new device starts the application and loads the offline Unity scene, the system automatically locates the image marker within the camera view and computes the offset while synchronizing the shared information.

We consider the following factors for aligning our coordinate system when loading a new scene. Specifically, the device world position is reset to $(0, 0, 0)$, while the rotation is only reset around the Y axis since each device can compute the correct X and Z rotations from the accelerometer. This requires us to compute the position and rotation of the camera differently in the shared coordinate space, shown in fig. 3. Let T_{cam} be the transform of the device camera in its local coordinate system when the image target is detected, and let T_{img} be the image target's transform in the local coordinate system. The

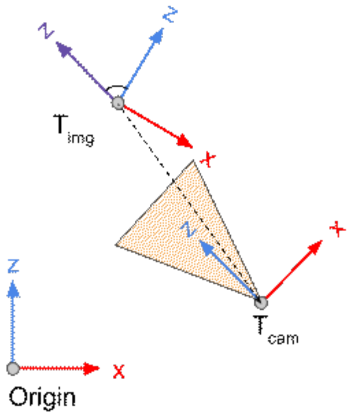

Fig. 3 Computing the initial camera offset in the common coordinate system. The position is the offset between the camera transform, T_{cam}, and the image target transform, T_{img}. The rotation is the angle between the projections of the Z vectors of these transforms on a horizontal plane.

position offset of the camera is computed as follows:

$$T_{offset} = T_{img}^{-1} T_{cam}$$

The position offset, T_{offset}, is the initial camera position in the shared coordinate system. To compute the initial camera rotation, we only use the rotation around the y axis. Let Z'_{cam}, Z'_{img} be the projections of the Z vectors of T_{cam} and T_{img} onto a horizontal plane, respectively. The rotation around the Y axis can be calculated by simply computing a quaternion q that rotates from Z'_{img} to Z'_{cam} as follows:

$$q_{xyz} = Z'_{img} \times Z'cam \tag{1}$$

$$q_w = Z'_{img} \cdot Z'_{cam} + \sqrt{\|Z'_{img}\|^2 \|Z'_{cam}\|^2} \tag{2}$$

Finally, we compute the absolute offset, which is relative to the image target's transform in the CAD model. Then, we multiply the relative offset by the corresponding image target's transform in the CAD model. This information is sent back to each device to translate between the individual and shared coordinate systems.

Synchronizing Locations in Real-Time. With the transformed locations in the shared coordinate system, we can continue to synchronize poses of individual users. We have implemented a custom network transform synchronization logic which utilizes a lower level constructs – syncvars. Specifically, we synchronize the local position and rotation independently. For the efficiency of the system, we only exchange information when the difference of pose is beyond a pre-assigned threshold. Once receiving the information, our server automatically synchronizes the changed syncvars across all clients. Then, each client

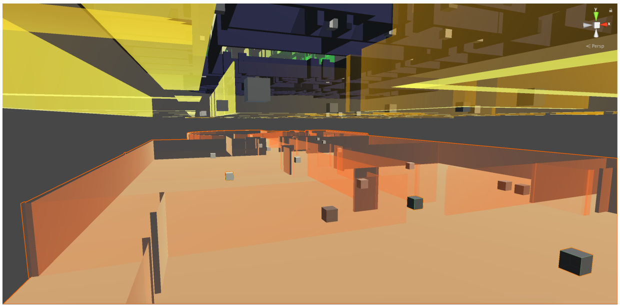

Fig. 4 We construct our own navigation system with key location points in a 3D Building structure.

transform the information into their own spaces. We also interpolate the position and rotation between received coordinates to achieve smooth animation. All objects or persons in our system are synchronized with this procedure.

4 Immersive Navigation, Visualization and Interaction

We present the three main components of our system for navigation, visualization and interaction functions respectively.

4.1 Immersive Building Navigation

We first build a navigation function to assist the exploration of large buildings. Similar to basic GPS navigation systems, we build our navigation with key node-points in the building structure. A node can represent a turn point, or a point of interest which can be an office, classroom, conference room, laboratory, or restroom etc. We model the nodes as empty GameObjects in Unity's hierarchy, which can be added to the model of the building directly (shown in Figure 4). The following lists several stages of our navigation system.

Node Graph Generation. We need to create a graph on the top of node-points to enable automatic navigation functions. This is achieved by composing a text file to denote the reachability from each node. For example, if there are three nodes in a building A, B and C, there will be three lines in the text file:
A C
B C
C A B
The first letter of each line denotes the starting node and the rest of the characters in the line tells if they are reachable from the starting node. Here C is reachable from A, and also B is reachable from C. A and B cannot reach each other directly.

Finding the Shortest Path. To represent the distances in the system, we can assign weights to the edges of the graph. Here we automatically assign the weights based on the 3D distances between the nodes. We can also assign weights to the nodes manually so that special locations can be emphasized. We then implement the breadth first search algorithm to find the shortest path from an assigned starting position to the destination.

Path Visualization. We then visualize the path on the mobile/AR devices by drawing the path through each node inside the selected pathway. While this can be implemented with Unity's Line Renderer component, we keep a global line renderer and reuse it for each new path drawn for efficiency. The path visualization is handled slightly differently between the two types of devices. During interactive exploration, since HoloLens needs to change its location based on user's position continuously, the Line Renderer does not stay intact with its parent component in Unity and hence may appear drifting away from the model as user moves away. To avoid this issue, we increase the frequency of redrawing the line renderer on HoloLens.

4.2 Immersive Building Visualization

We have designed different visualization to best suite user experience for different types of devices. The iOS devices mainly use the screen for 2D visualization and the HoloLens favor 3D visualization that can be mixed with the real world. The following presents four types of building visualizations, including 3 types for visualizing entire building structures and 1 type for AR based immersive navigation.

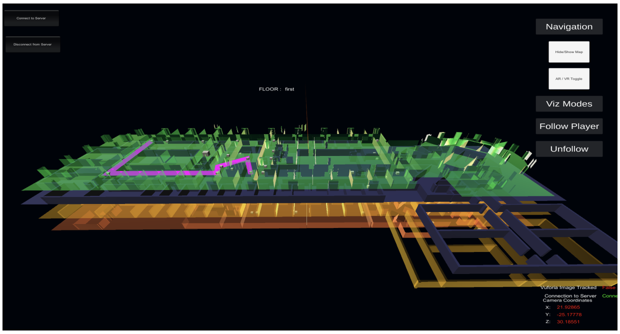

Fig. 5 Isometric building visualization on iPad shows the floor map and navigation path in 3D.

Isometric Visualization. We use isometric visualization [55] to represent the layered structure of a building created from the accurate physical model. As the default visualization method, our server transforms the coordinates of all users and converts them into other visualization types. We also render the floor model with a transparent shader material to allow a see through effect from any view angle.

The HoloLens allow us to visualize the 3D building model in the physical space. We implement a floating map, which re-orient the maps to face the user at all times. The maps also always followed the user and anchored itself to specific place in front of user. Figure 5 shows an example of isometric visualization on iOS device and the right of Figure 1 shows an example on HoloLens.

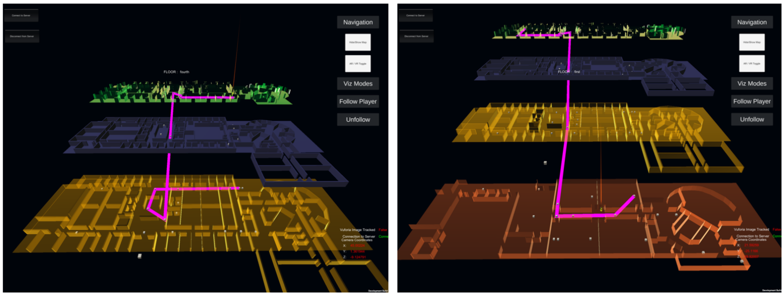

Fig. 6 Two examples of staircase visualization from different angels and navigation paths.

Staircase Visualization. The staircase Visualization is designed to avoid the overlapping issue of isometric visualization, while still allowing users to understand the 3D building structure. As shown in Figure 6, the layered staircase displays the original physical model in the open 3D space.

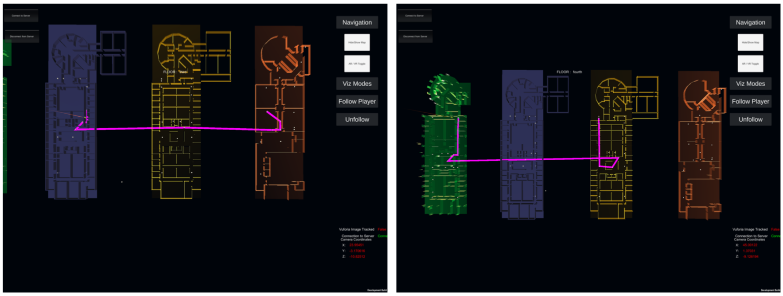

Fig. 7 Two examples of orthogonal visualization from different angels and navigation paths.

Orthogonal Visualization. The orthogonal Visualization [56] continues to flatten the 3D building structure side-to-side. As shown in Figure 7, this view avoids the occlusion issue in 3D building model completely while loses the 3D connections. It is suitable when users focus on one level of the model in the context of the entire building.

AR Immersive View. The AR immersive view is designed based on the participants' view. This information can be obtained from both mobile and AR devices. We use the user position and direction to control camera in the system, producing the first person view. This method produces the virtual building navigation effect and can be merged with the real physical environment during navigation. Figure 8 demonstrates examples of AR immersive view blended with real world.

Fig. 8 Two examples of AR immersive views on iPad, taken from different locations in a building. The right one also shows the navigation path in purple.

4.3 Immersive Interaction

We support our immersive system with several interaction functions. All of our interaction types support the touch controls on iOS devices and hand gestures and voice commands on HoloLens. To support immersive navigation, we also provide a world-in-minia- ture (WIM) views to improve situational awareness of participants. Our WIM shows building structures and real-time information about the users and important objects. Specifically, the follow methods are implemented to support smooth interactions.

Performance Improvement with Particle System. Simply displaying individual instances to visualize all information posts a significant performance bottleneck. To ensure the essential information is updated in real-time, we simplify the rendering of paths and recording targets only to positions. This allows the system to use a custom particle system containing all objects, instead of keeping individual instances, and it is significantly faster – achieving a speed of rendering hundreds of targets without a significant frame rate hit.

Tracking objects or participants on the WIM. To simplify tracking objects on the WIM, we develop a unity behavior to automate the process. The module uses three parameters: a reference to the model representing the object on the WIM, a boolean indicating whether a separate instance should be created or a particle system point, and a color for the particle system point. When the server generate an object on a client, this custom behavior adds the object to the WIMManager for tracking along with the appropriate parameters. We also allow clients to follow others by creating the shortest path to them and updating it in real time as they move around.

Hand Tracking. We use HoloLens's hand tracking to show and manipulate the WIM. When the user's hand is identified, we show the WIM hovering over the hand. Our hand tracker also supports the use of two hands to move or rotate the WIM. When the first hand is shown, the tracker saves its ID, activates the WIM, and moves it to the tracked position. Subsequent updates of the hand with that ID are used to update the WIM position. When a second hand appears (while the first hand is still tracked), the tracker saves its

ID and its initial position. Subsequent updates of the hand with the second ID are used to compute the angle of rotation to rotate the WIM.

Voice Commands. Voice commands are provided to support several basic interactions. They are relatively easy to use comparing to hand gestures. Specifically, "Stay" command is used to keep the WIM in a fixed physical location, while the "follow" command makes the WIM to move along the user. The "zoom in" and "zoom out" commands allow uses to adjust the sizes of the model.

Navigation Interaction. We implemented different approaches for

Fig. 9 User interface for selecting destination on an iOS Device, captured from iPhone simulator.

mobile and AR devices. For iOS devices, we create basic UI components to show a destination list, where it allow users to select starting and ending points. For HoloLens, we use the package of Holotoolkit to build the interface. With a custom Animator script, we keep the the destination list near the user, facing towards the user at all times. Figures 1 and 9 demonstrate how buttons are displayed on iOS device and HoloLens respectively. The building model is also controlled differently on the two types of devices. While our iOS system incorporates multi-touch interactions, the HoloLens system uses its full capability of two hand manipulation to move, rotate and scale the entire model of building per the client's choice.

5 Collaboration with Immersive Visualization

While the previous sections focus on the experiences of a single user, this section provides several immersive functions to support collaborations with immersive visualization. We also show a case study with two users collaborating in our academic building.

We target at collaborations between users in the fields and at the command centers, which are often required by emergency responders such as firefighters, police, and military soldiers. Currently, the communication between command centers and crews are often performed via voice, cameras, and possibly hand-held devices; which offer limited and in-efficient solutions. We develop

Springer Nature 2021 LaTeX template

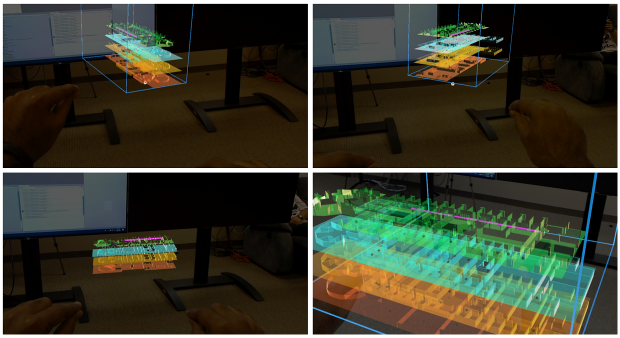

Fig. 10 Rotation and Scale using Two Hand Manipulation

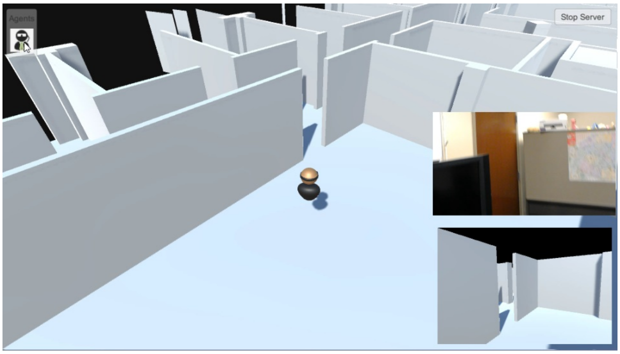

Fig. 11 The system for the command center visualizes all information collected in the system. The background snapshot shows a view from the server interface. Users can still wear AR devices at the center, and the two small snapshots on the right show the comparison of user view on the top and server view on the bottom.

the following collaborative functions to improve the coordination and situation awareness for multiple users performing real-time operations in smart buildings.

5.1 Approach

Server and Individual Views. Considering the different needs and situations of users, we design two types of system interfaces for users in the fields and users at the command center. The system for the command center can afford to visualize all information in the system, including all user locations, important marked targets, and the building structures. We allow the system to be viewed through computers at the command center, while also support immersive devices. The system for the users in the fields only shows important information including the targets and marked paths for real-time navigation. As shown in Figure 11, the users in the fields mainly see the real environment, while the server side shows the corresponding virtual environment.

The architecture of our platform provides a flexible mechanism to synchronize user locations in buildings and vital information across the team and command center in real-time. The workflow is demonstrated in Figure 12.

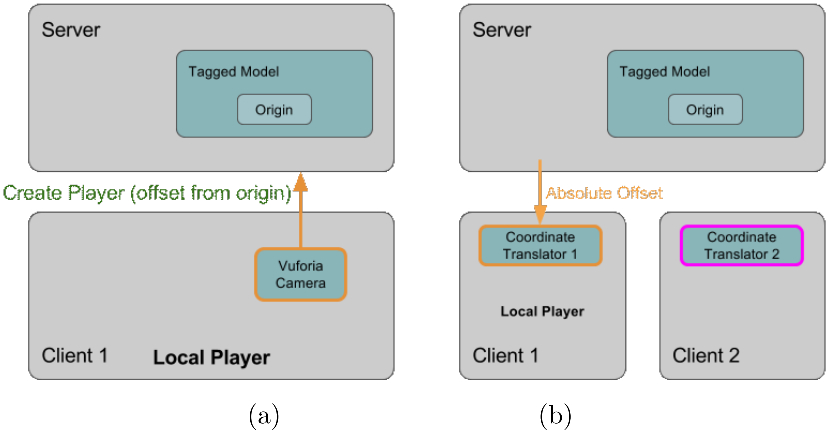

Fig. 12 Offset computation workflow. The workflow for computing the initial client transform in the common coordinate system. a) shows client1 before it connects to the server. The offline scene is loaded and the Vuforia camera is active. When the image target is located, a request is sent to the server to connect the client, and use the calculated offset. b) shows the server replying to client1 with its absolute offset, which is used to translate coordinates for that client. Notice that client2, which is already connected, has a different translator (translator2), since client2 has a different initial offset.

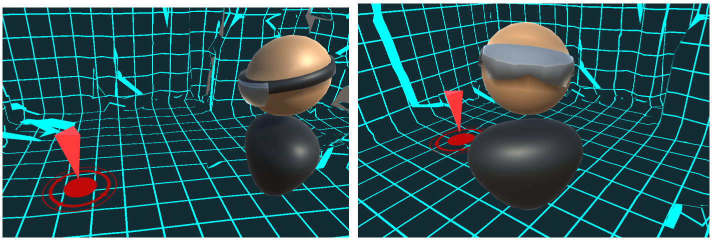

Fig. 13 The character representing operative locations in the field.

Avatar for Representing Crew Locations and Directions. To support understanding of the situations of crews in the field, we create avatars to represent individual users with a character that mimics the accurate head movement in real-time. The synchronized position and rotation are applied to the character head after interpolation, achieving smooth animation of the head. To make the characters look closer to a person instead of floating heads, we also add a body section which maintains an upright position while the head is moving. This is accomplished with a unity behavior attached to the body. The script has, as a parameter, an offset to determine how far below the head the body should be (to simulate a neck). The script moves the body to that location below the head, which only rotates vertically to face the same direction the user is looking. As shown in Figure 13, the result simulates character movement more naturally and is much more expressive than what only a floating head can be.

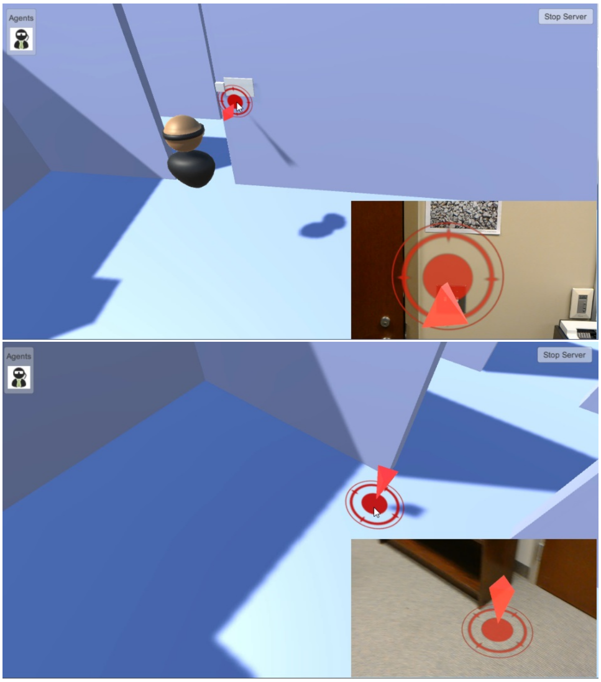

Fig. 14 Targets shown on the server and in mixed-reality. The server view is shown as the background and the HoloLens view is shown on the corner.

Creating Targets for Marking Important locations. We often need to mark important locations in a building, such as places with reported accidents. Targets can be created by dispatchers on the command server, or by users in the field, illustrated in Figure 14. In both cases, the server is responsible for creating and controlling the target instances. The dispatcher can left-click anywhere on the 3D CAD model to create a target. The target's location is synchronized across clients using the same custom network transform used for the characters.

For users to create targets, they issue a voice command ("Create Target"), which sends a command to create the target. The server raycasts the operatives head direction and creates a target at the hit point with the model.

Creating Paths for Detailed Instructions. To provide detailed instructions to the users, dispatchers at the command center can create paths on the server. This may represent safe paths for users to follow or paths with

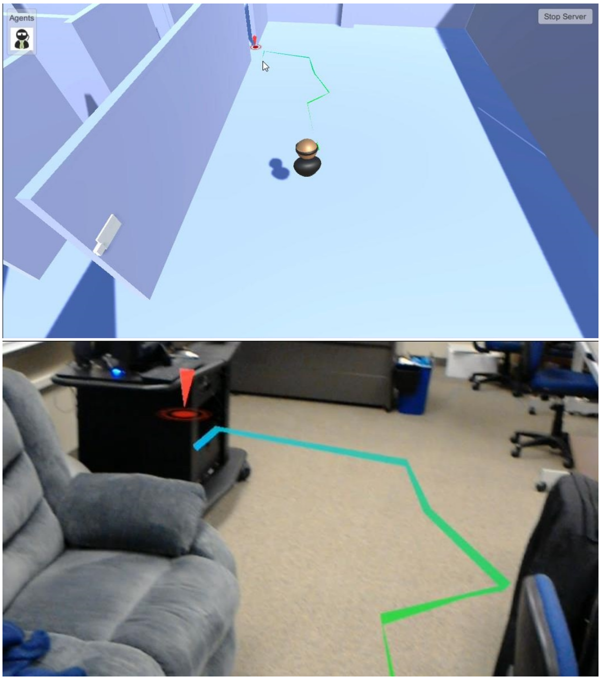

Fig. 15 Paths shown on the server and in mixed-reality.

all required locations to check. Right-clicking on the 3D CAD model initializes a path instance on the server (which is not synchronized yet). Subsequent left clicks add points to that path, until another right click is issued, which terminates the path and synchronizes it across the network clients (Figure 15).

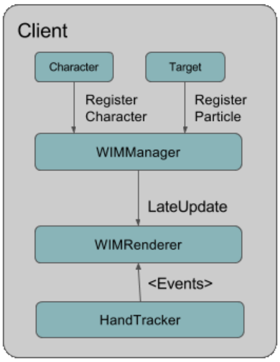

Fig. 16 General structure of the WIM system.

World in Miniature for On-Site Exploration. The system also supports a world-in-miniature (WIM) view to help improve users' situational awareness. We illustrate the general event structure of the WIM system in

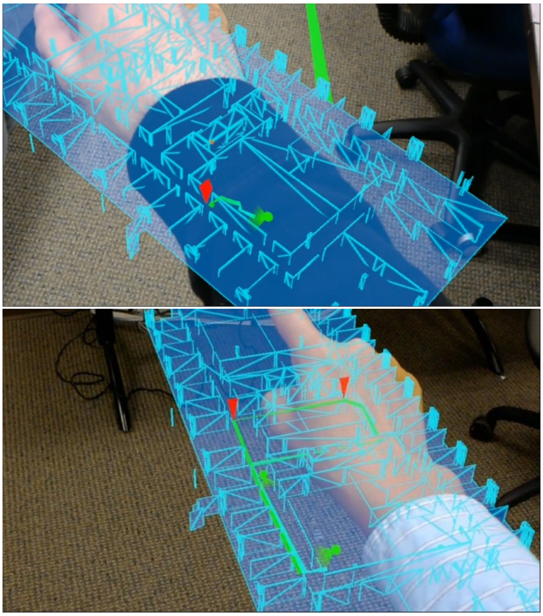

Fig. 17 World-in-miniature supports the users in the field to understand their environment and important locations of targets and other team members. The snapshot on the top shows the view for user in the field with only important locations, and the snapshot on the bottom shows the center view with locations of both users and more information than the user view.

Figure 16. Stoakley et al. explored the use of a WIM in VR [57], and concluded that WIMs offer many advantages and are intuitively used by users. As shown in Figure 17, our WIM shows real-time positions and rotations of crews in the field, as well as positions of all targets and paths. It utilizes HoloLens's hand tracking to show it hovering over the user's hand when it is within view. The WIM can be rotated using the other hand to view it from a better angle. Finally, the WIM allows users to create targets at locations that are not within their line of sight, by pointing their gaze point at the desired location and performing a tap gesture.

5.2 Case Study

We have tested our prototype system in one academic building on campus. Figure 18 demonstrates a sequence of interaction performed by a team of two users. As shown in the large background images, one user stays inside a room, holding the WIM with hand gesture and tracking the action of the other user. The second user follows the pre-assigned path (green to blue) and investigates the building. Wherever there is an issue, the second user can use hand gesture to mark the location either in the physical environment directly or on the WIM. The command center and both users can share all the information updated in real-time. Throughout such a long sequence of study, the second user has traversed around 50 meters in the building. All user locations, targets, and paths are synchronized in real-time accurately.

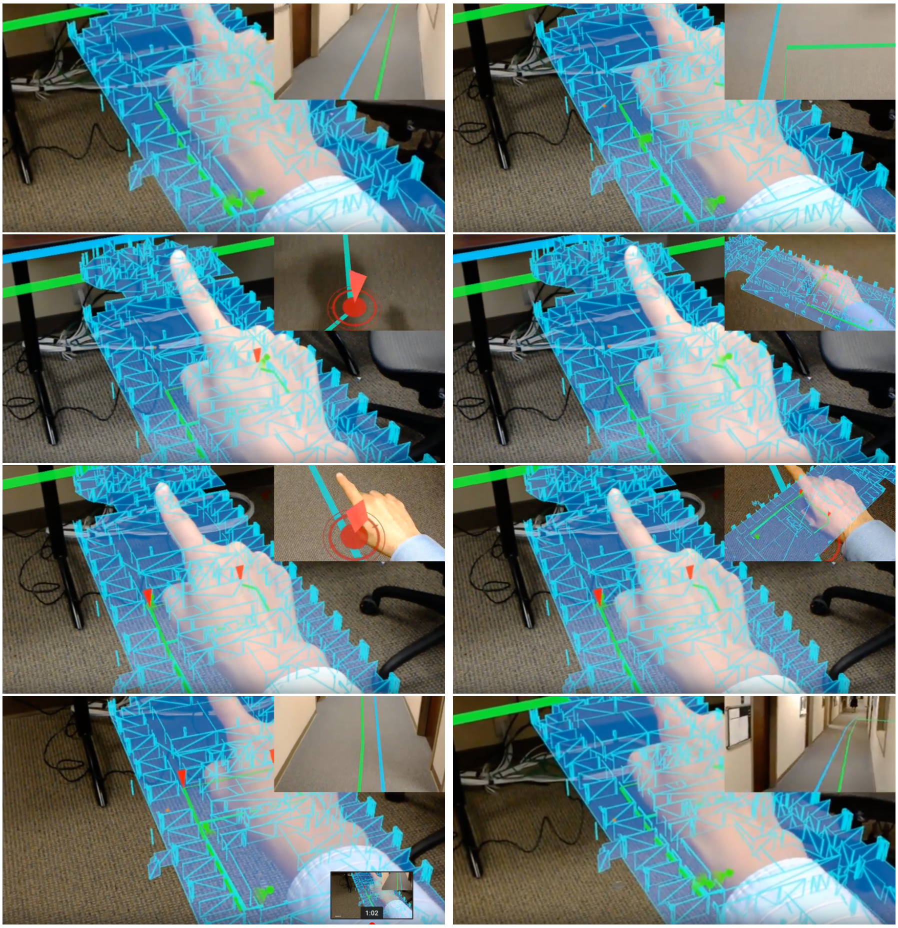

Fig. 18 Examples of a series of interactions in a case study, where two users (each wearing a HoloLens) coordinate with each other when exploring the building. The background images show the view from one HoloLens, and the embedded images show the view from the second HoloLens. The coordination and analysis between the users are achieved in real-time.

6 Evaluation and Discussions

We have conducted a user study to evaluate the effects of immersive navigation and acquired expert feedback. Since our interests are on the experience of using immersive navigation in real life scenarios, we have made two hypothesis: (1) immersive system improves the navigation tasks, (2) the performance of different participants may vary among different devices.

6.1 Evaluation

Procedure. To understand overall effectiveness of the system, we have designed four tasks involving different navigation functionalities.

Task 1: AR Immersive Nav System. The participants were asked to go to specific locations in the building. During the process, participants were asked to switch between immersive view and default view.

Task 2: Viz Exploration. The participants were asked to go to specific locations in the building via stairs. During the navigation, participants were asked to use any of the three visualization formats they feel comfortable with.

Task 3: Friend Finder. Group of two participants were considered in this task. Both participants registered the devices and then went to random locations inside the building. The task was to find each other out using our immersive system on iPhone or iPad. Participants were again asked to switch between any of the three visualization formats.

Task 4: Rescue Coordination. Group of two participants were considered in this task. One participant was asked to take iPhone or iPad, and the other participant was given a HoloLens. They both could talk to each other using cellular call. The two participants were asked to work together for one participant to reach a certain location on the map on fourth floor.

The tasks 1 and 2 were performed individually, where participants used either iPhone or iPad. In the tasks 3 and 4, we suggested participants to roam on one floor only to shorten the study time. The server kept track of the positions and rotations of all participants throughout the session.

Participants. Eight participants were recruited for user study. All of them were students of ages from 24 to 28. Two of them had visited this building before. No participants had experience of HoloLens previously, therefore they were given HoloLens to learn basic operations before starting the user study.

Expert Feedback. To acquire the psychological aspect of our approach, we consulted feedback from a researcher in the field of Psychological Science. She performed the same set of tasks in the study multiple times and tested all the mobile and AR devices. Her review states, "this is an innovative project with the potential to facilitate collaborative spatial tasks (e.g., navigation, search-and-rescue)." We also received valuable feedback on future improvements, including different colour schemes to distinguish users and improving the orientation of users.

Summary of Results. The tasks are designed from simple to complex. Among the feedback from participants, we find more positive feedback from complex tasks. The comments start from "could be useful but not required" in tasks 1/2 to "very useful" in tasks 3/4. Specifically, example comments from participants include – "Emergency Services like 911 to track down victim", "Finding items in big malls like Walmart or Lowe's or Costco", "Malls, Offices, parking areas" and "Shopping malls, Offices, multilevel parking lots."

6.2 Discussions and Limitations

The accuracy requirement for our system is not high, as it is navigation on the building scale. We have tested the system performance with both devices throughout our 4-floor academic building, which is shown in Figure 5. Our test includes different locations and paths from the same and different floors, including long paths from the first to the fourth floors or long paths traversing the hallways. Since both devices rely on environmental sensing with cameras, it is important to keep the environment consistent without sudden changes. There is generally no issue in the hallways, no matter the shapes such as straight paths or corners. However, getting into or out of the staircases requires attention. The sensors may lose track when the environment changes suddenly. Once lost, our system requires re-register the locations with any image marker stored in the system.

We find the accuracy of localization in our system satisfy the civilian need of indoor navigation. Accurate services for professional users may require strong signals and high-end localization inputs such as military level GPS. However, we do not require extra devices in our system. Our experience shows that our system with commercial devices is usable for civilian usages.

Our system is tested on the guest network, which offers only limited network access. Since we only exchange the locations and orientations of users and markers, the network bandwidth used in our system is very small. We are not aware of any obvious delay with our system.

The current system still need to be expanded significantly for professional usages, such as firefighters or police. When accuracy becomes more crucial, it is ideal to maintain the timing and network status in the system, so that users can make decisions based on the accurate information with the uncertainty degrees. We are aware that there are also many other visualization and interaction functions that can improve the communication efficiency and be added to our system. The security of the system should also be ensured for real-life usages.

7 CONCLUSION AND FUTURE WORK

This paper presents a collaborative immersive system which utilizes networked devices across-platform, including popularly used mobile devices (iOS devices) and AR HMD (HoloLens). Our approach for navigation improves users' communication in the physical building environments by supporting real-time visual communication on devices that can be carried freely. We have developed a set of building navigation, visualization, interaction and collaboration functions utilizing the available networking, iOS and HoloLens libraries. We have also performed a basic user study to evaluate the effects of immersive coordination and navigation. The results also suggest that such immersive functions are promising for challenging situations including complex building structures or time-critical navigation tasks. Our future work includes a plan to investigate additional methods to increase the capability of our navigation system.

We hope to demonstrate our approach for important applications of smart cities, especially for emergency responses such as firefighters and police.

References

[1] Azuma RT. A Survey of Augmented Reality. Presence: Teleoper Virtual Environ. 1997 Aug;6(4):355–385.

[2] Billinghurst M, Clark A, Lee G. A Survey of Augmented Reality. Found Trends Hum-Comput Interact. 2015 Mar;8(2-3):73–272.

[3] Kim K, Billinghurst M, Bruder G, Duh HB, Welch GF. Revisiting Trends in Augmented Reality Research: A Review of the 2nd Decade of ISMAR (2008–2017). IEEE Transactions on Visualization and Computer Graphics. 2018 Nov;24(11):2947–2962. https://doi.org/10.1109/TVCG.2018.2868591.

[4] Thangarajah A, Wu J, Madon B, Chowdhury AK. Vision-based registration for augmented reality-a short survey. In: 2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA); 2015. p. 463–468.

[5] Sutherland IE. A Head-mounted Three Dimensional Display. In: Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I. AFIPS '68 (Fall, part I). New York, NY, USA: ACM; 1968. p. 757–764. Available from: http://doi.acm.org/10.1145/1476589.1476686.

[6] Chen H, Lee AS, Swift M, Tang JC. 3D Collaboration Method over HoloLens^{TM} and Skype^{TM} End Points. In: Proceedings of the 3rd International Workshop on Immersive Media Experiences. ACM; 2015. p. 27–30.

[7] Roux WL. THE USE OF AUGMENTED REALITY IN COMMAND AND CONTROL SITUATION AWARENESS. Scientia Militaria - South African Journal of Military Studies. 2011;38(1).

[8] Santos MEC, Chen A, Taketomi T, Yamamoto G, Miyazaki J, Kato H. Augmented Reality Learning Experiences: Survey of Prototype Design and Evaluation. IEEE Transactions on Learning Technologies. 2014;7(1):38–56.

[9] Roberts JC, Ritsos PD, Badam SK, Brodbeck D, Kennedy J, Elmqvist N. Visualization beyond the Desktop–the Next Big Thing. IEEE Computer Graphics and Applications. 2014;34(6):26–34.

[10] Ayyanchira A, Mahfoud E, Wang W, Lu A. Cross-Platform Immersive Visualization and Navigation with Augmented Reality. In: Proceedings of

the International Symposium on Visual Information Communication and Interaction. VINCI; 2021. .

[11] Feiner S, MacIntyre B, Hollerer T, Webster A. A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment. In: Proceedings of the 1st IEEE International Symposium on Wearable Computers. ISWC '97. Washington, DC, USA: IEEE Computer Society; 1997. p. 74–81. Available from: http://dl.acm.org/citation.cfm?id=851036.856454.

[12] Tache R, Abeykoon HA, Karunanayaka KT, Kumarasinghe JP, Roth G, Fernando ONN, et al. Command Center: Authoring tool to supervise augmented reality session. In: 2012 IEEE Virtual Reality Workshops (VRW). IEEE; 2012. p. 99–100.

[13] Huang W, Sun M, Li S. A 3D GIS-based interactive registration mechanism for outdoor augmented reality system. Expert Systems with Applications. 2016;55:48 – 58. https://doi.org/http://dx.doi.org/10.1016/j.eswa.2016.01.037.

[14] Irizarry J, Gheisari M, Williams G, Walker BN. InfoSPOT: A mobile Augmented Reality method for accessing building information through a situation awareness approach. Automation in Construction. 2013;33:11 – 23. Augmented Reality in Architecture, Engineering, and Construction. https://doi.org/http://dx.doi.org/10.1016/j.autcon.2012.09.002.

[15] Gerstweiler G. Guiding People in Complex Indoor Environments Using Augmented Reality. In: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR); 2018. p. 801–802.

[16] Sayapogu T, Dsa K, Kaul P. AR Smart Navigation System. In: 2021 2nd International Conference for Emerging Technology (INCET); 2021. p. 1–4.

[17] Pillette L, Moreau G, Normand JM, Perrier M, Lecuyer A, Cogne M. A Systematic Review of Navigation Assistance Systems for People with Dementia. IEEE Transactions on Visualization and Computer Graphics. 2022;p. 1–1. https://doi.org/10.1109/TVCG.2022.3141383.

[18] Trapp M, Dumke F, Döllner J. Occlusion Management Techniques for the Visualization of Transportation Networks in Virtual 3D City Models. In: Proceedings of the 12th International Symposium on Visual Information Communication and Interaction. VINCI'2019. New York, NY, USA: Association for Computing Machinery; 2019. Available from: https://doi.org/10.1145/3356422.3356445.

[19] Perea P, Morand D, Nigay L. [POSTER] Halo3D: A Technique for Visualizing Off-Screen Points of Interest in Mobile Augmented Reality. In: 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct); 2017. p. 170–175.

[20] Gustafson S, Baudisch P, Gutwin C, Irani P. Wedge: Clutter-Free Visualization of off-Screen Locations. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '08. New York, NY, USA: Association for Computing Machinery; 2008. p. 787–796. Available from: https://doi.org/10.1145/1357054.1357179.

[21] Baudisch P, Rosenholtz R. Halo: A Technique for Visualizing off-Screen Objects. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '03. New York, NY, USA: Association for Computing Machinery; 2003. p. 481–488. Available from: https://doi.org/10.1145/642611.642695.

[22] Gruenefeld U, Ali AE, Boll S, Heuten W. Beyond Halo and Wedge: Visualizing out-of-View Objects on Head-Mounted Virtual and Augmented Reality Devices. In: Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '18. New York, NY, USA: Association for Computing Machinery; 2018. Available from: https://doi.org/10.1145/3229434.3229438.

[23] Wu ML, Popescu V. Efficient VR and AR Navigation Through Multiperspective Occlusion Management. IEEE Transactions on Visualization and Computer Graphics. 2018;24(12):3069–3080. https://doi.org/10.1109/TVCG.2017.2778249.

[24] Wang L, Wu W, Zhou Z, Popescu V. View Splicing for Effective VR Collaboration. In: 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR); 2020. p. 509–519.

[25] Wang L, Wu J, Yang X, Popescu V. VR Exploration Assistance through Automatic Occlusion Removal. IEEE Transactions on Visualization and Computer Graphics. 2019;25(5):2083–2092. https://doi.org/10.1109/TVCG.2019.2898782.

[26] Sukan M, Elvezio C, Oda O, Feiner S, Tversky B. ParaFrustum: Visualization Techniques for Guiding a User to a Constrained Set of Viewing Positions and Orientations. In: Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology. UIST '14. New York, NY, USA: Association for Computing Machinery; 2014. p. 331–340. Available from: https://doi.org/10.1145/2642918.2647417.

[27] Lee Y, Choi W, Sohn BS. Immersive gesture interfaces for 3D map navigation in HMD-based virtual environments. In: 2018 International

Conference on Information Networking (ICOIN); 2018. p. 963–965.

[28] Nam JW, McCullough K, Tveite J, Espinosa MM, Perry CH, Wilson BT, et al. Worlds-in-Wedges: Combining Worlds-in-Miniature and Portals to Support Comparative Immersive Visualization of Forestry Data. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR); 2019. p. 747–755.

[29] Coffey D, Malbraaten N, Le TB, Borazjani I, Sotiropoulos F, Erdman AG, et al. Interactive Slice WIM: Navigating and Interrogating Volume Data Sets Using a Multisurface, Multitouch VR Interface. IEEE Transactions on Visualization and Computer Graphics. 2012;18(10):1614–1626. https://doi.org/10.1109/TVCG.2011.283.

[30] Biocca F, Tang A, Owen C, Xiao F. Attention Funnel: Omnidirectional 3D Cursor for Mobile Augmented Reality Platforms. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '06. New York, NY, USA: Association for Computing Machinery; 2006. p. 1115–1122. Available from: https://doi.org/10.1145/1124772.1124939.

[31] Tonnis M, Klinker G. Effective control of a car driver's attention for visual and acoustic guidance towards the direction of imminent dangers. In: 2006 IEEE/ACM International Symposium on Mixed and Augmented Reality; 2006. p. 13–22.

[32] Wither J, White S, Azuma R. Comparing spatial understanding between touch-based and AR-style interaction. In: 2011 10th IEEE International Symposium on Mixed and Augmented Reality; 2011. p. 273–274.

[33] Mulloni A, Seichter H, Schmalstieg D. User experiences with augmented reality aided navigation on phones. In: 2011 10th IEEE International Symposium on Mixed and Augmented Reality; 2011. p. 229–230.

[34] Faria NdO, Gabbard JL, Smith M. Place in the World or Place on the Screen? Investigating the Effects of Augmented Reality Head -Up Display User Interfaces on Drivers' Spatial Knowledge Acquisition and Glance Behavior. In: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW); 2020. p. 762–763.

[35] Zollmann S, Hoppe C, Langlotz T, Reitmayr G. FlyAR: Augmented Reality Supported Micro Aerial Vehicle Navigation. IEEE Transactions on Visualization and Computer Graphics. 2014;20(4):560–568. https://doi.org/10.1109/TVCG.2014.24.

[36] Hein P, Bernhagen M, Bullinger AC. Two is Better Than One. Improved Attention Guiding in AR by Combining Techniques. IEEE Computer Graphics and Applications. 2020;40(5):57–66. https://doi.org/10.1109/

MCG.2020.3012274.

[37] Morar A, Băluţoiu MA, Moldoveanu A, Moldoveanu F, Butean A, Asavei V. Evaluation of the ARCore Indoor Localization Technology. In: 2020 19th RoEduNet Conference: Networking in Education and Research (RoEduNet); 2020. p. 1–5.

[38] Khenak N, Vézien J, Bourdot P. Spatial Presence, Performance, and Behavior between Real, Remote, and Virtual Immersive Environments. IEEE Transactions on Visualization and Computer Graphics. 2020;26(12):3467–3478. https://doi.org/10.1109/TVCG.2020.3023574.

[39] Pankratz F, Dippon A, Coskun T, Klinker G. User awareness of tracking uncertainties in AR navigation scenarios. In: 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR); 2013. p. 285–286.

[40] Kishishita N, Kiyokawa K, Orlosky J, Mashita T, Takemura H, Kruijff E. Analysing the effects of a wide field of view augmented reality display on search performance in divided attention tasks. In: 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR); 2014. p. 177–186.

[41] Kim K, Billinghurst M, Bruder G, Been-Lirn Duh H, F Welch G. Revisiting Trends in Augmented Reality Research: A Review of the 2nd Decade of ISMAR (2008–2017). IEEE Transactions on Visualization and Computer Graphics. 2018 09;PP:1–1. https://doi.org/10.1109/TVCG.2018.2868591.

[42] Sereno M, Wang X, Besancon L, Mcguffin MJ, Isenberg T. Collaborative Work in Augmented Reality: A Survey. IEEE Transactions on Visualization and Computer Graphics. 2020;p. 1–1. https://doi.org/10.1109/TVCG.2020.3032761.

[43] Szalavári Z, Schmalstieg D, Fuhrmann A, Gervautz M. "Studierstube": An environment for collaboration in augmented reality. Virtual Reality. 1998;3(1):37–48. https://doi.org/10.1007/BF01409796.

[44] Benko H, Ishak EW, Feiner S. Collaborative mixed reality visualization of an archaeological excavation. In: Third IEEE and ACM International Symposium on Mixed and Augmented Reality; 2004. p. 132–140.

[45] Nilsson S, Johansson BJE, Jonsson A. Cross-Organizational Collaboration Supported by Augmented Reality. IEEE Transactions on Visualization and Computer Graphics. 2011;17(10):1380–1392. https://doi.org/10.1109/TVCG.2010.249.

[46] Dong S, Kamat VR. Collaborative visualization of simulated processes using tabletop fiducial augmented reality. In: Proceedings of the 2011 Winter Simulation Conference (WSC); 2011. p. 828–837.

[47] Butscher S, Hubenschmid S, Müller J, Fuchs J, Reiterer H. Clusters, Trends, and Outliers: How Immersive Technologies Can Facilitate the Collaborative Analysis of Multidimensional Data. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. CHI '18; 2018. p. 90:1–90:12.

[48] Elmqvist N, Irani P. Ubiquitous Analytics: Interacting with Big Data Anywhere, Anytime. IEEE Computer. 2013;46(4):86–89.

[49] Kister U, Klamka K, Tominski C, Dachselt R. GraSp: Combining Spatially-aware Mobile Devices and a Display Wall for Graph Visualization and Interaction. Comput Graph Forum. 2017 Jun;36(3):503–514. https://doi.org/10.1111/cgf.13206.

[50] Horak T, Badam SK, Elmqvist N, Dachselt R. When David Meets Goliath: Combining Smartwatches with a Large Vertical Display for Visual Data Exploration. In: Proceedings of the ACM Conference on Human Factors in Computing Systems; 2018. Available from: http://www.umiacs.umd.edu/~elm/projects/david-goliath/david-goliath.pdf.

[51] Speicher M, Hall BD, Yu A, Zhang B, Zhang H, Nebeling J, et al. XD-AR: Challenges and Opportunities in Cross-Device Augmented Reality Application Development. Proc ACM Hum-Comput Interact. 2018 Jun;2(EICS):7:1–7:24. https://doi.org/10.1145/3229089.

[52] Cui Z, Sen S, Badam SK, Elmqvist N. VisHive: Supporting Web-based Visualization through Ad-hoc Computational Clusters of Mobile Devices. Information Visualization. 2018;.

[53] Weick KE. The Mann Gulch disaster - The collapse of sensemaking in organizations [Journal Article]. Administrative Science Quarterly. 1993;38(4):628–652.

[54] Weick KE. South Canyon Revisited - Lessons from High Reliability Organizations [Journal Article]. Wildfire. 1995;4:54–68.

[55] Ben-Haim D, Lappan G, Houang RT. Visualizing rectangular solids made of small cubes: Analyzing and effecting students' performance. Educational Studies in Mathematics. 1985 Nov;16(4):389–409. https://doi.org/10.1007/BF00417194.

[56] Wang X, Lim J, T Collins R, Hanson A. Automated Texture Extraction from Multiple Images to Support Site Model Refinement and Visualization. In: The Fourth Int. Conf. in Central Europe on Computer Graphics and Visualization; 1996. .

[57] Stoakley R, Conway MJ, Pausch R. Virtual Reality on a WIM: Interactive Worlds in Miniature. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '95. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co.; 1995. p. 265–272. Available from: http://dx.doi.org/10.1145/223904.223938.