Computers & Graphics 84 (2019) 24–41

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Technical Section

# Towards large scale high fidelity collaborative augmented reality

Damien Constantine Rompapas[a,*], Christian Sandor[b], Alexander Plopski[a], Daniel Saakes[c], Joongi Shin[c], Takafumi Taketomi[a], Hirokazu Kato[a]

[a] Interactive Media Design Lab, Nara Institute of Science and Technology, Nara, Japan
[b] eXtended Reality Lab (XRL), School of Creative Media, City University of Hong Kong, Kowloon Tong, Hong Kong, China
[c] Department of Industrial Design, KAIST, Daejeon, South Korea

A R T I C L E I N F O

*Article history:*
Received 14 May 2019
Revised 23 August 2019
Accepted 24 August 2019
Available online 27 August 2019

*Keywords:*
Augmented reality
High fidelity
Game design
Large scale interaction

A B S T R A C T

In recent years, there has been an increasing amount of Collaborative Augmented Reality (CAR) experiences, classifiable by the deployed scale and the fidelity of the experience. In this paper, we create HoloRoyale, the first large scale high fidelity (LSHF) CAR experience. We do this by first exploring the LSHF CAR design space, drawing on technical implementations and design aspects from AR and video games. We then create and implement a software architecture that improves the accuracy of synchronized poses between multiple users. Finally, we apply our target experience and technical implementation to the explored design space. A core design component of HoloRoyale is the use of visual repellers as crowd control elements to guide players away from undesired areas. To evaluate the effectiveness of the employed visual repellers in a LSHF CAR context we conducted a user study, deploying HoloRoyale in a 12.500 $m^2$ area. The results from the user study suggest that visual repellers are effective crowd control elements that do not significantly impact the user's overall immersion. Overall our main contribution is the exploration of a design space, discussing several means to address the challenges of LSHF CAR, the creation of a system capable of LSHF CAR interactions along with an experience that has been fitted to the design space, and an indepth study that verifies a key design aspect for LSHF CAR. As such, our work is the first to explore the domain of LSHF CAR and provides insight into designing experiences in other AR domains.



## 1. Introduction

Augmented Reality (AR) superimposes computer graphics (CG) onto the user's view of the real world. In recent years, there has been an increasing amount of Collaborative Augmented Reality (CAR) experiences [1–3], classifiable by the deployed scale (from 16 $m^2$ room areas to urban areas greater than 10,000 $m^2$) and the fidelity of the experience. Kruijff et al. [4] identified several issues that affect the fidelity of an AR experience. From the identified issues, we describe the fidelity of an AR experience by the following metrics:

- *Virtual-real interactions:* Does a virtual object behave like its real world counterpart? This includes physical collisions and occlusions. An example would be a virtual ball interacting with a real wall.
  *High fidelity:* A real wall in front of a virtual ball will occlude it. When the ball is thrown at the wall, it will bounce off of the wall.
  *Low fidelity:* The virtual ball will always be visible, contradicting the depth placement of the two objects. When thrown, the ball will pass through the real wall.
- *Accurate content registration:* Is the placement of virtual content consistent with the real world context? An example is a virtual statue being placed on a bust.
  *High Fidelity:* If correctly registered the statue will appear on the bust.
  *Low Fidelity:* A bad registration will lead to the statue visually floating in the air.
- *Spatio-temporal consistency:* When an interaction occurs, do all users see the action at the same time and place? For example, a user throwing a virtual rock.
  *High Fidelity:* When there is spatio-temporal consistency all users observe the virtual rock being thrown at the same time, with the origin of the thrown rock at the users hand.

* Corresponding author at: Interactive Media Design Lab, Nara Institute of Science and Technology, Nara, Japan.

*E-mail addresses:* hyperlethalvector92@gmail.com, damien.rompapas.dk1@is.naist.jp (D.C. Rompapas), chris.sandor@gmail.com (C. Sandor), plopski@is.naist.jp (A. Plopski), saakes@kaist.ac.kr (D. Saakes), jgshin@kaist.ac.kr (J. Shin), taketomi.takafumi@ieee.org (T. Taketomi), kato@is.naist.jp (H. Kato).

https://doi.org/10.1016/j.cag.2019.08.007
0097-8493/

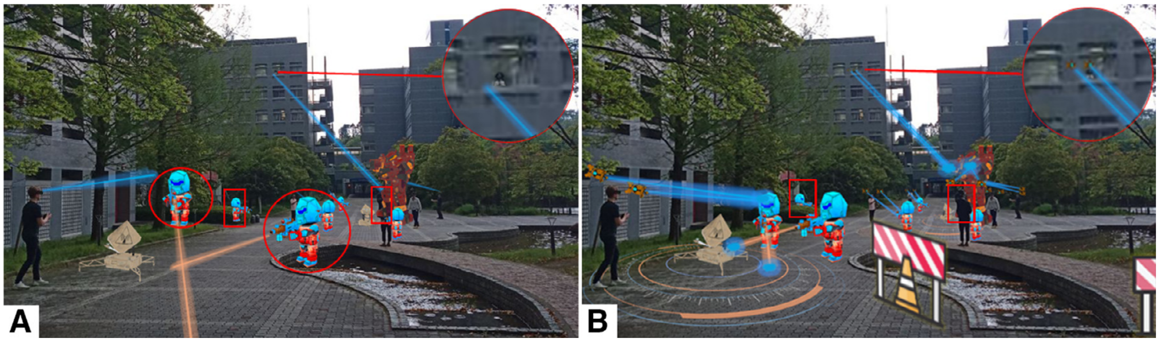


**Fig. 1.** The aim of this work is to create and deploy a Collaborative Augmented Reality (CAR) experience on a university sized area, with high fidelity (HF) AR content. (A) Current large scale (LS) CAR experiences only exhibit low fidelity content such as: inaccurate registrations (zoomed section, red circles), missing occlusions (red squares) and no interactions between the real and virtual environment. (B) We achieve accurate content registration over large distances (zoomed section), correct occlusions (red squares) and interactions between the real and virtual environments. We hide spatial and temporal inconsistencies by representing users as remote avatars (drones). We also incorporate the following user redirection elements: attractors (highlighted satellites) to guide users towards key locations and repellers (roadwork signs) to keep users away from areas that are dangerous/prone to system failure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

*Low Fidelity:* Spatial inconsistencies cause the virtual rock to appear to be thrown from a visually incorrect origin, or hit an incorrect target. Temporal inconsistencies cause delays between the user's movement and the virtual rock being thrown.

- *Visual consistency:* Does a virtual object have the same visual properties as the surrounding environment and their real world counterparts. This can be considered from two aspects, geometrical quality and lighting consistency. An example is a virtual copy of a real statue being shown side by side.
  *High fidelity:* The virtual object is indistinguishable from its real world counterpart as the geometry of the virtual object is dense and able to represent the smooth surface of the real world statue accurately, and the lighting condition of the virtual environment is accurate to the real world.
  *Low fidelity:* The virtual object can be easily identified by either a mismatch in lighting, or by deformations in the geometrical model used to render the virtual object.

Although visual consistency plays an important role in achieving perfect fidelity, it presents significant challenges such as accurate estimation of the illumination and the scene reflectance [5–7], transparency [8], realistic rendering of shadows [9], and replication of other visual effects exhibited when observing real objects, such as the depth of field [10]. While an experience that does not replicate photorealistic elements can still be high fidelity, if users are aware that non-photorealistic rendering is justified by story and artistic elements. Conversly, a photorealistic experience will not necessarily be high fidelity, for example if it has significant temporal inconsistencies. As such, we focus on the fidelity issues related to interactions.

Experiences that target large areas [2,3,11], commonly have only rudimentary interactions with the physical world, suffer from content registration errors, or exhibit spatio-temporal inconsistencies and therefore do not cover many of the fidelity issues described by [4]. We classify these experiences as large scale and low fidelity (LSLF). On the other hand, various room sized experiences [12–14] satisfy all of the fidelity metrics. We classify these experiences as RS and high fidelity (RSHF). Although it's technically possible to track multiple users with high accuracy in a large scale environment using Simultaneous Localization and Mapping (SLAM) [15], there are several technical limitations, such as the accuracy of synchronized poses between users and network latency.

Our goal is to create the first LSHF CAR experience by addressing these challenges. In particular, we aim to create a multiplayer AR game deployed in a suburban area larger than 10,000 $m^2$, featuring high fidelity content (Fig. 1). To achieve this, we have to not only address the technical challenges, but also consider additional design issues unique to LSHF CAR. We organized a week long workshop between Institute 1[1] and Institute 2[1], seven researchers from HCI, Augmented Reality, Game and Industrial Design backgrounds gathered together to discuss these design issues. We reviewed prior implementations [3,16–19] and identified the following core design issues:

- Users will be moving over a large area, and can potentially move into hazardous areas (such as a busy road or a staircase) or areas that the utilized system may not function within (dark areas) [20].
- Interactions will occur over large areas, within a single contiguous instance. Additionally, the actions of one user will affect the global state (An example is a sniper taking a virtual shot over a long distance, or a user activates a virtual button at one location, triggering a door in a separate location to open).
- Users will be distributed over the large area and will need an understanding of their environment, the situation within the experience, and the intentions of non co-located users (An example is a team of users working together at separate locations to achieve a goal).
- The input device used to interact with the virtual content can exhibit errors in tracking, making the interactions difficult, especially over large distances [11].

Then, through affinity diagramming [21] we grouped these challenges into the following four clusters:

- *User redirection*: How to move users around the play area, directing them towards key locations (attractors) and away from dangerous/unplayable areas (repellers).
- *Inconsistencies:* How to handle spatio-temporal inconsistencies during runtime, providing a consistent experience for all users.
- *Spatial awareness:* How to provide the users with information about the surrounding real/virtual environment, and the location of other users.
- *Communication:* How to provide a means of communicating between non co-located users.

Although AR research has extensively explored communication and navigation in LS CAR environments [20,22,23], user redirection and handling spatio-temporal inconsistencies have yet to be addressed. We can adapt game design elements to specifically address these design issues as they share the same design issues [24]. Ng et al. [25] utilized video games elements to navigate users within a room scale environment. Although, they did not consider the use of game elements outside the game context, they highlight the necessity of user redirection elements.

In this paper, we derive the requirements needed to achieve our target LSHF CAR experience. Based on these requirements we explore the LSHF CAR design space, drawing on technical implementations and design aspects from both AR and video games. We then present a software architecture and technical implementation that improves the accuracy of synchronized poses between multiple tracking systems. We apply our target experience and technical implementation to our established design space, creating HoloRoyale, the first instance of a LSHF CAR experience. One of the most pressing concerns we identified during the workshop is keeping users away from potentially hazardous areas or areas that the system cannot be used in. Because of this, a core design component of HoloRoyale is the use of visual repellers to guide players away from dangerous/unplayable areas. While demonstrating HoloRoyale in smaller scale demonstration venues, we found that users became frustrated with the placement of the repellers, but respected their boundaries. This raised the question if this was due to the scale of the venues and what other effects repellers could have on users immersed into a LSHF CAR. This prompted us to evaluate the effectiveness of the employed visual repellers in a LSHF CAR context. To do this, we conducted a user study, deploying HoloRoyale in a 12.500 $m^2$ area. The results confirm that visual repellers are effective user redirection elements that do not significantly impact the user's overall immersion. Our results also show that peer and time pressure can lead to users ignoring repellers, which requires their effects to be reinforced by means of additional design elements. Furthermore, we found that repellers complicate communication between users as they make it more difficult to maintain a mental image of the environment layout.

Overall, we make the following contributions:

1. Our work is the first to explore the challenges of LSHF CAR.
2. We establish a design space that offers a new approach and perspective to handle the requirements of LSHF CAR experiences by adapting concepts from video game design.
3. We improve the accuracy of synchronized poses between multiple SLAM systems by aligning several smaller SLAM maps, creating a global coordinate system. We track each user relative to the smaller SLAM maps, avoiding pose drift over large areas. Our framework enables the creation of future LSHF CAR experiences on a global scale.
4. We create the first instance of a LSHF CAR experience by applying our technical implementation and our target LSHF CAR experience to our established design space.
5. The results from our evaluation show that virtual repellers can be effective user redirection elements in LSHF CAR contexts. This leads to new research questions on the benefit of user redirection elements and how to reinforce the effect they provide.

The work in this paper is a first step into the previously unexplored domain of LSHF CAR, opening up several new avenues for future work. Besides investigating the effects of adapted game design elements on users in LSHF CAR scenarios, there are several research questions that can now be investigated. What other AR spaces can benefit from the adaptation of game design into AR? How does hiding spatio-temporal inconsistencies impact the performance of users in LSHF CAR scenarios? What are the psychological impacts of diegetic repellers when represented as dangerous obstacles?

## 2. Requirements

Our aim is to create a high fidelity AR experience that is deployed on a university/suburban sized scale with multiple simultaneous users. As our target experience covers the fidelity challenges we derived from [4] and the unique LSHF CAR challenges ascertained during the workshop described in Section 1, we can consider it an experience that encompasses all challenges expected in a LSHF CAR experience. Through affinity diagramming [21] we categorize these challenges as requirements based on the component of the experience that they affect. The result is the following list of general requirements for a LSHF CAR experience (Fig. 2):

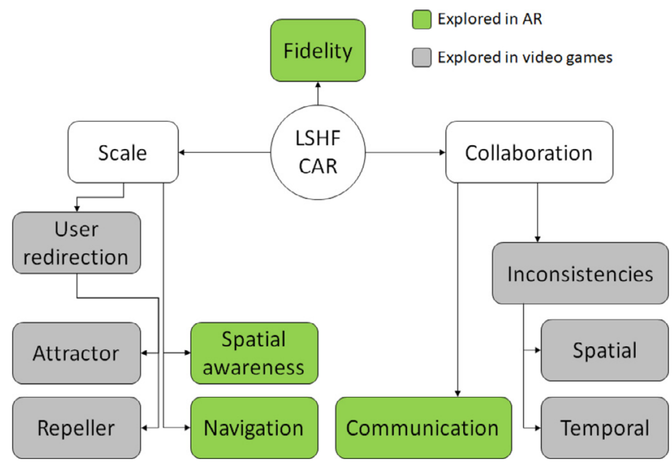


**Fig. 2.** Taxonomy summarizing the dependency of system requirements for LSHF CAR applications and the domains that have explored methods to address these requirements.

Scale

- The system must be deployable in areas up to and beyond a maximum size of 10,000 $m^2$, to cover the target university sized area.
- Due to users moving around a larger area, the system must be able provide users with information about their surrounding environment.
- To assist users' movements over the larger area, the system must provide navigation cues to assist players when moving between key locations.
- Since it is expected that user's encounter dangerous situations, move into areas where the system may no longer work, or be unaware of the next destination, the system must provide the following user redirection elements:
  - Repellers to deter users from entering dangerous/unplayable areas.
  - Attractors to highlight key locations, motivating users to move towards them.

Fidelity

- To provide visually realistic CG the system must:
  - Render high density geometry
  - Accurately model the lighting conditions of the real environment
- The system must produce a 3D model of the environment for virtual-real interactions and visual occlusions.
- The display and input must have a total motion-to-photon latency no larger than 20 ms to prevent motion sickness while moving around the large area [26].
- The system must be able to render convincing geometrical models of virtual objects, whenever applicable.
- To ensure that the virtual content appears consistent within the environment its displacement in the user's view must be less than 1 arcmin [27].

Collaboration

- To enable a collaborative environment, the system must share the pose and logical state of several clients.
- To provide a consistent experience between clients, the system must handle:
  - Inconsistent between-client temporal states.
  - Erroneous between-client pose synchronization.
- To support collaboration between users, the system must provide a means of communication between users.

Although we listed a generic set of requirements for LSHF CAR experiences, we do not address the fidelity issues related to visual consistency as it is beyond the scope of this work. However, as we aim to cover the majority of the fidelity requirements, we still consider our target experience one of high fidelity.

## 3. Design space

In this section, we discuss how these requirements can be addressed from a technological and a user interface standpoints. We first explore two key areas for the technical platform: how to display content to the user, and how to track the user in the real world. By categorizing related work by the technical implementation used, the fidelity it achieves, and the scale of deployment, we can identify the most suitable display and tracking technology for our target experience. Then we explore several aspects of the user interface and user experience. Hereby, we gather insights not only from previous AR implementations, but also from different genres in video games. Video games are widely imagined in AR [28] and have to handle many of the design challenges present in LSHF CAR. Fig. 2 shows the design requirements for our target LSHF CAR experience and the domains that have previously explored them. A summary of our established design space resulting from the discussion in this section can be seen in Fig. 3.

### 3.1. Technical platform

Various aspects of our requirements, such as display latency and tracking accuracy, can be addressed by selecting the appropriate platform. In particular, we consider the following areas: How to display content to the user, and how to track the user within the environment.

#### 3.1.1. Displays

One crucial component of any AR system is displaying CG content to the user's view of the real world. In general, there are three ways to show AR content to users:

*Video See-Through (VST):* This method composites CG content onto a video stream. This technology is commonly used in HMDs [3,29–32] and hand-held devices [2].

*Optical See-Through (OST):* This method directly embeds CG into the user's view of the environment by reflecting a rendering from a screen off a transparent half-mirror, into the user's eye. Several experiences [11–13,20,23,33,34] utilize OST-HMDs.

*Projection based:* This method projects CG directly onto the environment. Although often used in AR [35], projector are typically statically placed, and are limited to presenting content onto the physical world (which limits the depth perception of CG content). Therefore, projection based AR is not viable for our scenario. Our goal is to deploy our LSHF CAR experience in a suburban area. Therefore, we need to consider that users will be moving between indoor and outdoor areas. In this scenario, a display exhibiting a motion-to-photon latency larger than 20ms [26] can lead to dangerous situations such as users walking into an object or falling over due to motion sickness. There have been extensive comparisons of OST and VST-HMDs [36] that suggest VST displays are more restricted on the motion-to-photon latency. This is because users view the world through the video camera feed, with the CG composited, as opposed to OST-HMDs that directly render the CG content onto the user's view of the environment. Additionally, when a VST-HMD fails, users can no longer see their surroundings. This explains why most LS AR experiences utilize OST-HMDs [11,20,23]. OST-HMDs however rely on a half mirror to present content to the user, under bright lighting conditions the external light transmission causes the exhibited CG to appear more transparent, affecting the fidelity of the content shown. While it is possible to address this through brighter displays and occlusion-capable systems [37], currently no commercially available system provides this functionality. VST-HMDs and hand-held VST displays do not suffer under this condition as the video camera feed can adjust it's exposure to match the surrounding environment light.

Overall OST-HMDs are the best candidate for our LSHF CAR experience, as they satisfy the motion-to-photon latency requirement and are fail-safe. Additionally, hand-held VST can be used for non-immersive AR experiences [38].

#### 3.1.2. Tracking

To place AR content and synchronize the poses of several users, we must obtain each user's pose in the environment. For this, there are three main approaches:

*Sensor based tracking:* Uses the GPS, accelerometer, gyroscopic sensor, and compass on the device to obtain the position and orientation of the user, within the real world [2,3,11,20,23]. Although easily accessible, sensors are prone to drift and inaccurate readings [39], which can cause severe content registration issues [11]. These sensors do not rely on any visual input for tracking and are therefore robust to differences in lighting conditions.

*Outside-in tracking:* Obtains the user's pose by utilizing external sensors placed within the environment. A common approach is to track fiducial markers placed on the user [29,30]. Although these systems can achieve high accuracy, the setup becomes excessively expensive when deploying over larger areas and requires careful calibration and preparation. Additionally, sunlight can negatively affect the tracking accuracy. However, since outside in tracking does not require natural features of the environment (and instead typically relies on retro-reflective markers that reflect IR light emitted from the mounted sensors) it performs very well under low light conditions.

*Inside-out tracking:* This method functions similar to outside-in tracking. However, the sensors (most commonly cameras) are placed on the user and track features within the environment. These features can be either fiducial markers placed throughout the scene [32] or natural features [15].

Although it is possible to use markers for LS environments [11], this requires careful between-marker calibration [40]. Furthermore, the user's pose can only be estimated when a marker is detected by the sensors.

The alternative utilizes natural features detected within the camera image to localize and track a user in the environment (Simultaneous Localization and Mapping, SLAM [15]). Recent improvements enable the use of SLAM on mobile systems [41] and track users even over large scales [42]. Nevertheless, pose drift occurs when tracking the user over large areas, even when using loop closure to minimize this error [42,43]. These inside-out tracking methods are more robust in daylight scenarios but fail under low light conditions due to the lack of trackable features in the environment.

*Hybrid:* This method combines different tracking methods to leverage their advantages. RSHF experiences such as those shown on the Microsoft HoloLens [12] and the Magic Leap [13] utilize multiple carefully calibrated cameras for visual SLAM, depth

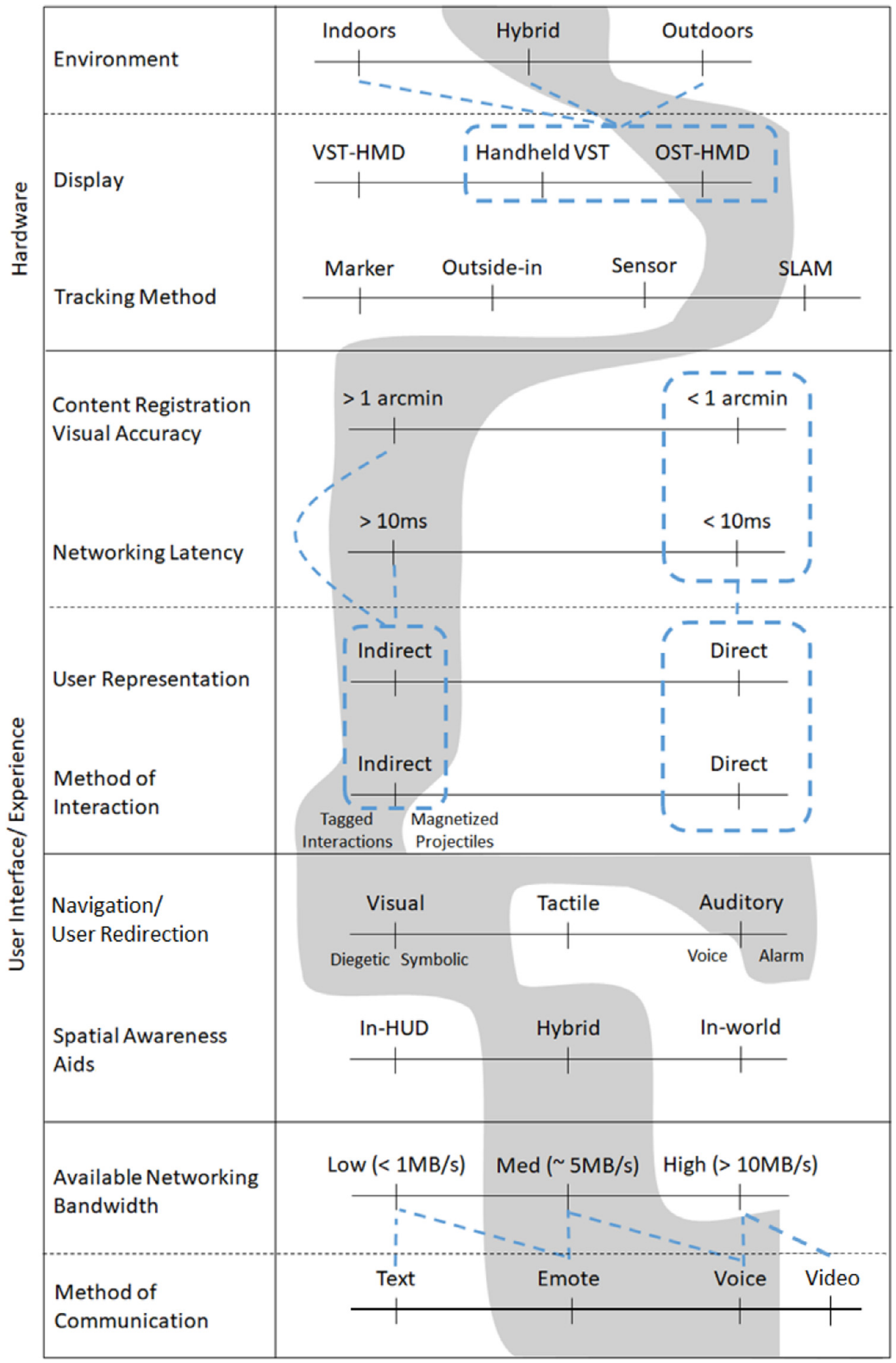


**Fig. 3.** A morphological chart showing our established design space. The blue dotted lines indicate our general guidelines based on the discussion in Section 3. The grayed out area represents how our technical implementation and our target experience fit within our established design space. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

sensors for surface mapping, as well as gyroscopes and accelerometers to improve the tracking stability. Nevertheless, these systems also suffer the same pose drift issues over large scales. As this method still relies on some form of visual SLAM, it suffers from the same issues under low light conditions.

We assume that our target experience will only be played during daylight hours. With this assumption, although hybrid methods still suffer from pose drift issues, their improved accuracy and off-the-shelf availability makes them the prime candidate for our target LSHF CAR experience.

### 3.2. User interface and experience

Some of the requirements listed in Section 2 require careful design of the user interface and the CAR experience.

*3.2.1. Representing the user in the AR environment*

Many CAR experiences assume that users are in perfect sync both spatially and temporally [1]. However, the nature of distributed experiences means that spatial and temporal inconsistencies are present due to tracking errors and network latency. These inconsistencies can severely disrupt the fidelity of an experience. We can hide possible spatial and temporal inconsistencies by modifying how we represent the user in AR. We can represent users in the AR environment through:

*Direct representation:* This representation is used by most AR applications. It utilizes the raw pose of tracked users and tools when placing CG into the scene. Although this is the ideal scenario, it is only viable if there are no spatial or temporal inconsistencies. An example of direct representation is the rendering of a gun over the controller in the user's hand [13].

*Indirect representation:* This method represents the user/tool as an unattached AR avatar. It, therefore, overcomes spatial inconsistencies by disassociating the user from the virtual environment. Furthermore, by interpolating and predicting the pose and state of the associated avatar [44], indirect representation hides temporal inconsistencies. For example, virtual wands could represent users in a magic game [45].

Since our target LSHF CAR experience features multiple distributed users, we expect temporal inconsistencies to occur. This favors indirectly representing users in the AR environment.

*3.2.2. Interacting with content in CAR*

We need to consider how users interact with content, as this is a core component of any AR experience.

For any interaction to occur, users must first select a target for interaction. Fitts' Law [46] states that the time taken to select a target is determined by the distance from the user to the target and the size of the target. The shorter the distance and the larger the size, the easier it is to point at the target. Spatio-temporal inconsistencies vary the effective width and distance of a target increasing the difficulty of selection. Overall, users can interact with content in the following ways: Direct interaction: Direct interaction with virtual content appears to be most natural and is applied in a variety of AR experiences [1,11–13]. However, as this interaction utilizes the user's raw input, it is highly susceptible to tracking errors, inaccurate pose synchronization, and network latency. Under such conditions, direct interaction can result in reduced efficiency and increased player frustration [11].

*Assisted interaction:* Similar to direct interaction, assisted interaction uses the user's raw input for interaction. However, it improves the robustness to spatio-temporal inconsistencies by modifying the effective width and distance of a target without modifying its visual appearance. This allows interaction to occur, even if it's not visually consistent but can cause frustration if the assistance constantly selects wrong target [47].

*Indirect interaction:* This interaction technique is widely used in video games [48–50] and is only possible when users and their tools are represented by avatars. Hereby, the avatars always orient themselves towards the interaction target selected by users and perform the desired interaction. This interaction method can be further enhanced by applying assisted interaction techniques to the user's input for target selection. Indirect interaction is robust to pose synchronization errors and network latency, providing consistent interactions [51] while allowing incorrect selections if the user's aim is imprecise.

*Magnetized interaction:* This method is specific to projectiles. Hereby, the projectile acts like a 'heat seeking missile' continuously changing its flight path as it moves towards the intended target, independent of the user's input. Although it ensures consistent interactions, this method removes all challenge from the experience and can lead to dissatisfaction [47].

As we determined that we should indirectly represent users to overcome temporal inconsistencies, indirect interaction is most suited as it preserves spatial and temporal consistency. Although magnetized interaction could address this issue as well, it removes the challenge from the experience.

*3.2.3. Communication between users*

When users are distributed over large areas they need a means of communication with each other. From the grouping of related work, and examining the communication methods utilized in many video games, we identify four generally used types of communication:

*Text based communication:* Users communicate by sending a string of characters typed out on a virtual on-screen keyboard or a physical input device [52]. This method provides clear communication and requires minimal networking bandwidth. However, creating and reading a message is time consuming and causes an increased cognitive load [53]. Therefore, it should be avoided if possible.

*Emoticon based communication:* Instead of typing out messages, users can utilize a predetermined set of emoticon text messages or images based on the user's possible intentions. Emoticon-based communication is widely used in video games [54]. Emoticon messages have the benefit of being fast to send, are instantly understandable (requiring minimal cognitive load) and utilize minimal network bandwidth. However, due to the limited range of options, the intention that a user wants to portray can often be ambiguous.

*Voice based communication:* As an alternative to visual communication, many collaborative experiences use voice chat [1]. This offloads communication from visual to auditory, decreasing cognitive load [53]. It has also been shown to be preferred to text based communication in collaborative environments [53]. The drawback of voice based communication is the high networking bandwidth demand.

*Video based communication:* Instead of communicating only over voice, several games feature video based communication where users see either a first-person view or a view of the partner's face during communication [55]. Although such communication is often used in collaborative systems [1], it requires a much higher per-user networking bandwidth. Furthermore, it may not significantly improve the collaboration due to the limited size of the shared view and difficulties understanding what their partners mean when many users share their view at the same time. This makes it difficult to use in scenarios with more than 2–3 concurrent users.

To our knowledge, there has been no study in a distributed AR context that directly compares video, voice, text, and emoticon based communication between users. Nevertheless, voice based communication is the prime candidate for our target LSHF CAR experience as it allows clear and fast communication. However, if the available bandwidth is not sufficient enough to support voice based communication, we can add emoticon based communication as an alternative.

*3.2.4. Providing spatial awareness*

As users are exploring a large environment with AR content added to it, they require a method to obtain an understanding of their surroundings. This includes information related to the task, the environment, and the location of users. There are three ways to provide this information.

*2D representations in the Heads Up Display:* This method places spatial awareness cues into the 2D plane that lies in screen space (also known as the Heads Up Display or HUD). This representation can contain varying degrees of detail ranging from simple radars [56] to detailed maps of the environment [20]. The HUD can also contain cues for out-of-view points of interest [57]. However, adding too many elements to the HUD can also lead to visual clutter of the display [58].

*3D representations in the user's environment:* This method uses a 3D model representation of the environment, displayed within the user's viewport (instead of in screen space). Such a world in miniature (WIM) [59] shows the users' location within their environment [60] and any additional contextual information [61]. This technique is also commonplace in video games as both a symbolic and diegetic element embedded into the game environment [62]. The downside to this representation is that in order to provide detail it has to occupy a large portion of the screen space, potentially occluding the user's view of the environment.

*Hybrid representations:* Finally, there are hybrid implementations that combine both 2D and 3D representations of the environment. For example, when the user is looking at an AR scene it can be annotated by 2D labels shown on the HUD. Then, when the user views the WIM, the labels move to their corresponding positions on the WIM [63]. Although hybrid representations retain the benefits of both 2D and 3D representations they require careful consideration on when the content should switch between its 2D to 3D representations.

Although all of the listed methods are viable, we utilize a hybrid representation as it is the most powerful of the three.

#### *3.2.5. Navigation and User Redirection in AR:*

When exploring large scale areas, navigation and user redirection cues become necessary. These elements help direct users towards intended areas and lead them away from areas that are hazardous or prone to system failure. These elements also function as navigation aids. There are two key types of user redirection elements:

*Attractors:* These elements highlight areas of interest, prompting users to move towards them.

*Repellers:* These elements highlight areas where users are not allowed to enter by either indicating danger, or inaccessibility.

Visual user redirection elements can appear as symbolic elements in the HUD [22,23]. For example, an icon flashing on the screen is an attractor while a text prompt warning users if they enter an unwanted area is a repeller.

Instead of relying purely on symbolic in HUD elements, video games also employ diegetic user redirection elements to maintain the experience's immersion [64]. For example, a signal flare in the distance or a robot guide are both diegetic attractors. On the other hand, burning walls of fire or a closed door act as diegetic repellers by indicating that the blocked section is either dangerous or inaccessible. Ng et al. [25] utilized diegetic video games elements to navigate users within a RS game environment. However, they did not consider their use as user redirection elements outside the game context.

There are also several non visual cues usable for user direction. Audible voice feedback directly conveys necessary information to users, but can distract users from their current task [65]. Audible alarms are another alternative, but are vague if there's no context for the alarm [66]. Finally we can consider vibro-tactile feedback that has been shown to be effective at navigating users with vision deficiency [67]. However, these cues are also vague without a given context.

Since our LSHF CAR experience targets a suburban environment, clear representation of navigation and user redirection elements is key. Although visual redirection elements have been shown to be most effective in similar environments within video games [68] it is unclear how effective they will be in LSHF CAR as the virtual object rendered on an AR display will not physically prevent users from entering the repellers bounds, their visibility may be obstructed by other elements in the environment, or users may plainly be distracted by other pedestrians and the immersive gameplay. At the same time, symbolic cues could be more obvious. This suggests that a combination of different cues should be used to overcome the limitations of each system.

## 4. Creating a system capable of LSHF CAR

Although combination of SLAM and sensor-based tracking offers the best approach for tracking users in large scale environments, system drift can lead to severe errors when sharing user poses (Fig. 7c). In this section, we describe a client server architecture that improves the accuracy of synchronized poses between multiple users over large distances (Fig. 7d).

### *4.1. Hardware selection*

From the technical analysis in Section 3, we find that currently the Microsoft HoloLens and Magic Leap One are the ideal hardware to deploy our experience on. Both devices feature compelling RSHF experiences [12,13]. At the time of development, the Magic Leap was not commercially available, in consequence, we built our LSHF CAR system around the Microsoft HoloLens. The HoloLens is an OST-HMD with a motion-to-photon latency of less than 20 ms [69] and has a microphone built into it, allowing voice communication. The Microsoft HoloLens contains a sensor assisted SLAM system for tracking the user and provides a 3D reconstruction of the surrounding environment that can be used for near-distance occlusions and virtual-real environment interactions. However, the tracking system inside the HoloLens can experience pose drift over large distances, limiting its deployable scale in CAR. Additionally the HoloLens has a limited FOV for augmented content, limiting the fidelity of the experience by deteriorating the visual consistency [4,36]. Nevertheless, we opted for the HoloLens as our target platform as it addresses many of the requirements of our system and presents a fail-safe platform. The next subsection details how we extend the usable range of the HoloLens to satisfy the scale requirements for LSHF CAR.

### *4.2. Software architecture*

To satisfy our scale requirements for LSHF CAR, we must synchronize the pose of multiple users within a LS area. For this we have two options, using the HoloLens poses directly (which are susceptible to drift) or synchronizing via a cloud computed global coordinate system. There are already cloud solutions for localizing and sharing the pose of several clients in a single collaborative environment such as 6D.ai [70] and immersal [71], however, it is unclear if these cloud solutions support our scale requirements. Furthermore, they are incompatible with the HoloLens. Instead of these cloud based solutions, we extend [72], taking several smaller mapped areas, but additionally computing transformations between the maps. This allows the poses of all clients and virtual actors to be synchronized into a single global coordinate system. We propose a client-server based architecture that performs the following steps to create a global coordinate system (Note that for the purposes of adaptability, we describe the design and implementation in abstract terms applicable to any Visual SLAM system, and mention the relevant HoloLens specific implementation terms in brackets):

*Preparation*: During preparation, we scan several areas up to $100m^2$ using the Microsoft HoloLens. An origin of each mapped area is tagged (a HoloLens anchor is placed in the scene), and the 3D model, along with the tagged origin and binary data (HoloLens anchor data) that represents the VSLAM map, is uploaded to the alignment server. The alignment server then creates a global map, computing the transforms between each map origin (HoloLens anchor) by performing a series of bounding-box Iterative Closest

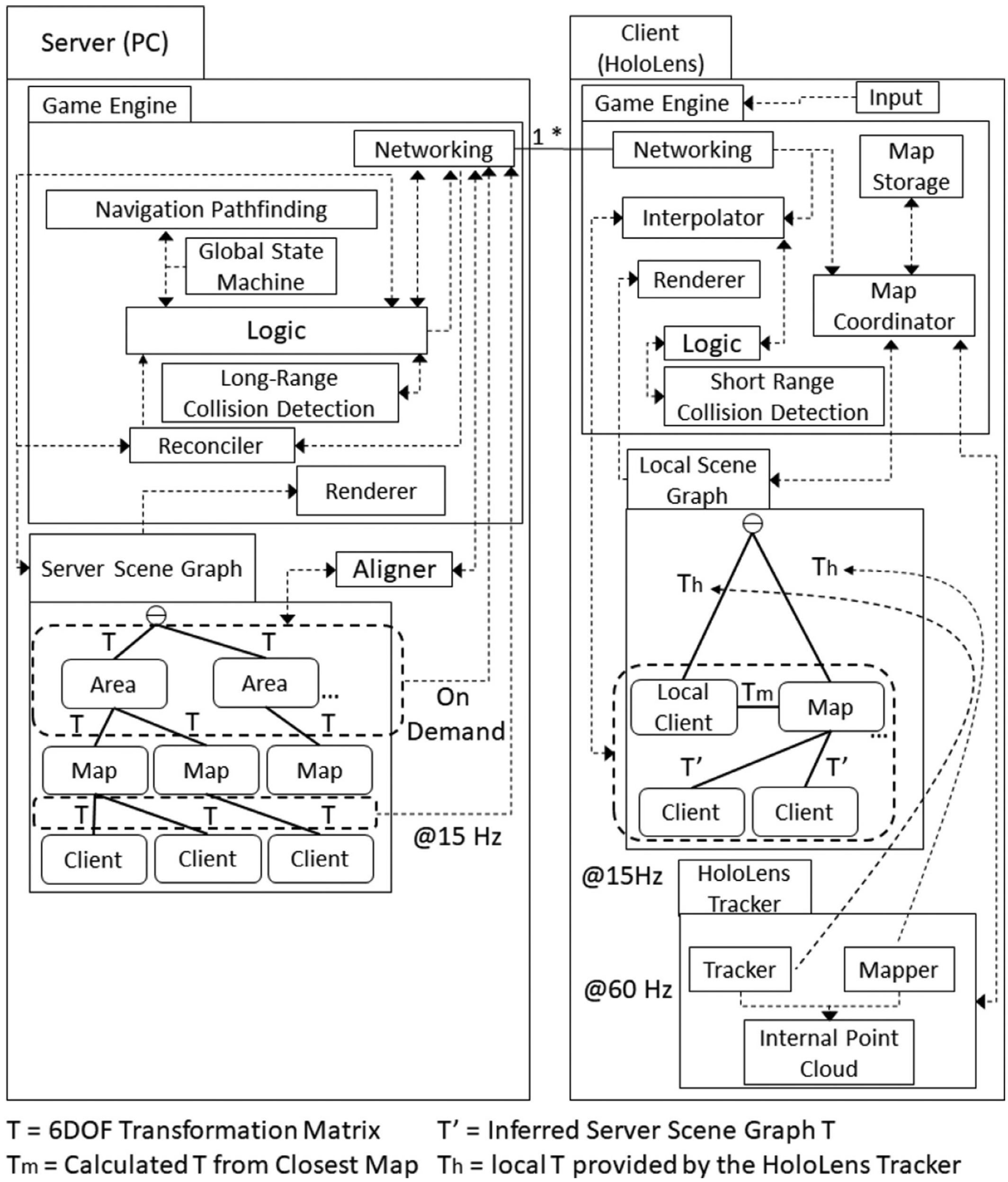


**Fig. 4.** System decomposition with the scene graph sections expanded. The 6DOF transform matrices between maps are static, synchronized on demand. The client transforms relative to their closest tracked map are synchronized at 15 Hz.

Point (ICP) [73] alignments using the 3D models. This is done by attempting an alignment for each pairwise model along each side of a 6 sided cube, then accepting the alignment that contains the minimal amount of error, and saving the transformation for that alignment. The completed scene graph is stored in a database for later use. We later author AR content directly onto the aligned global 3D model. The pose of the content is computed relative to the closest map origin.

*Distribute poses:* On system start-up we assume the HoloLenses start in an assumed starting location and query a web api with this location. The web api streams several candidate maps (HoloLens anchors) to the HoloLens. The maps (HoloLens anchors) are sequentially loaded into the HoloLens' internal tracker until it localizes a loaded map (Places the anchor into the scene). Once localized, we track the HoloLens relative to the localized map's origin (HoloLens anchor), sending the relative transform to a game server that then computes it's pose in respect to the global scene graph.

We then distribute the resulting updated global scene graph to all clients. To minimize networking bandwidth, the distribution is done in two parts. The static between-map (HoloLens anchor) transforms from the alignment are sent on demand. The computed poses of each HoloLens and computer-controlled virtual actors are synchronized at 15 Hz. The poses are interpolated between frames as described in [44] (See Fig. 6a). The 15Hz synchronization rate for poses can be extended up to 60 Hz to provide higher precision, at the cost of an increased networking load.

A system decomposition that outlines the timing for sending subsections of the global scene graph can be seen in Fig. 4. As the HoloLens moves through the area mapped out during preparation, additional maps (HoloLens anchors) are loaded and localized (placed into the scene). The HoloLens is always tracked relative to the closest localized map origin (HoloLens anchor). If a map cannot be localized, it is flagged on the server. Once a map is flagged by three separate clients, the origin (HoloLens anchor) is removed from the scene graph (with the 3D model retained). Then, a new

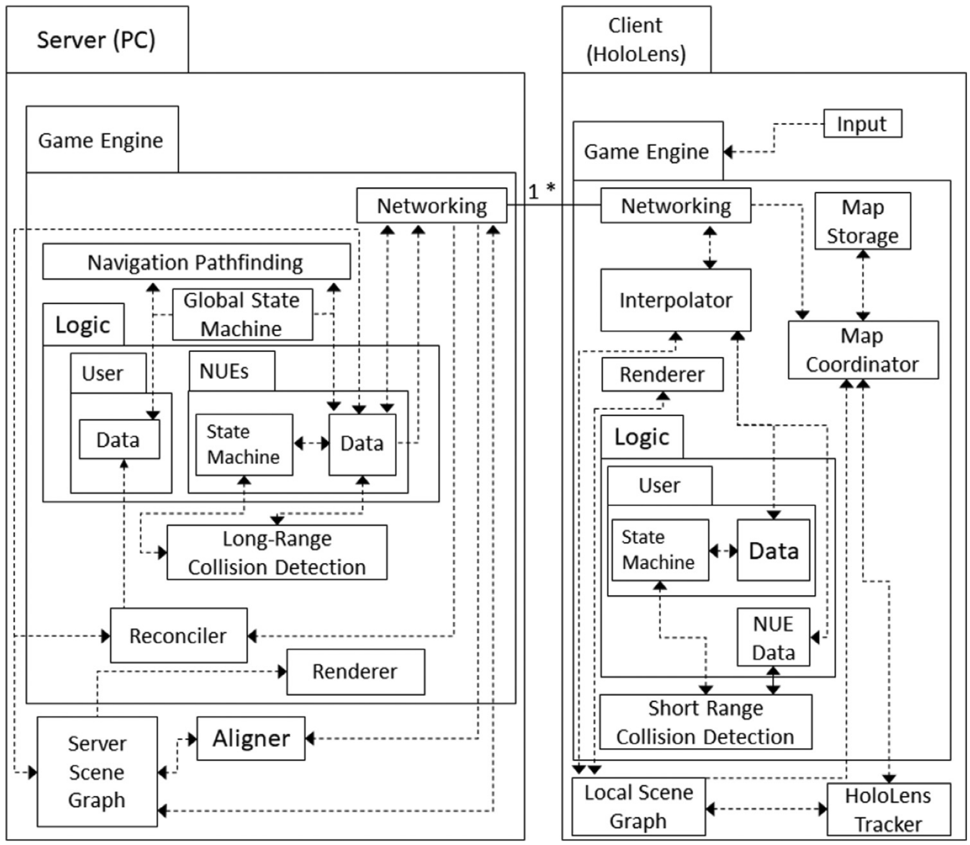


**Fig. 5.** System decomposition with the game engine components expanded. This shows how the majority of computation and logic is offloaded onto the server. The server handles all of the experiences logic, including the state machines for all Non User Entities (NUEs). The client only runs a minimal viewer, processing the state of the local client, and short range collision detection.

map origin (HoloLens anchor) and 3D model of the surrounding area of a nearby client is captured and uploaded.

*Place AR content:* AR content is placed according to the global scene graph. During runtime, as a rendering optimization we use Load on Demand to switch the model used for visual occlusions and interactions. We use a manually prepared cube-based phantom model of the environment at distances larger than 5 meters, and the HoloLens spatial mapping model at distances shorter than 5 meters (the effective range of the depth sensor).

*Offloading computation to the server:* For computations we run a game engine on both the client and the server. During runtime we offload the majority of computation to the server. The client runs a minimized viewer, only interpolating the current local state based on incoming state updates from the server. The server processes the non user entities and transmits the states to the clients. To optimize collision detection, we utilize a combination of short-ranged collision detection on clients (as they contain the most recent model of the environment) and long-ranged collision detection on the server (as it contains a global map and can perform a higher rate of collision detection without impacting performance). The results of all interactions are reconciled on the server (See Fig. 6c). A complete system decomposition with a focus on the offloaded core components can be seen in Fig. 5.

### 4.3. Implementation

The following describes the specific hardware the system was implemented on and the software that the system was developed with:

*Client:*

The client runs on the Microsoft HoloLens, utilizing an XBox One S Controller for input and Mobile Wi-Fi networking. The software consists of the Unity game engine (2018.3.1f1) that comprises of C++ and C# code.

*Alignment & Game Server:*

Although it is possible to run the alignment and game servers on separate machines, we deploy both on a single Microsoft Surface Book 2 laptop computer with the following specs:

- Intel Quad-Core i7-8650U, @ 4.2 GHz
- RAM: 16 GB DDR4
- GPU: Nvidia GeForce GTX1060, 6 GB

The alignment server utilizes a RESTful web api developed on Golang (9.2) and uses a PostgreSQL database to store static poses. The game server is built using the Unity game engine (2018.3.1f1) that comprises of C++ and C# code.

### 4.4. Visual verification

We performed a visual verification to test the accuracy of AR content placement in screen space and pose synchronization using our system against using an out of the box HoloLens for pose synchronization. We placed two HoloLenses with infrared LEDs attached running our system in a previously mapped and aligned environment. Then augment the view from each HoloLens with a colored virtual crosshair placed according to the pose resulting from the synchronization system used (Red = our system, Blue = native HoloLens). We then oriented both HoloLenses so that they face each other roughly 5, 25, 50, and 75 m apart. We compared the

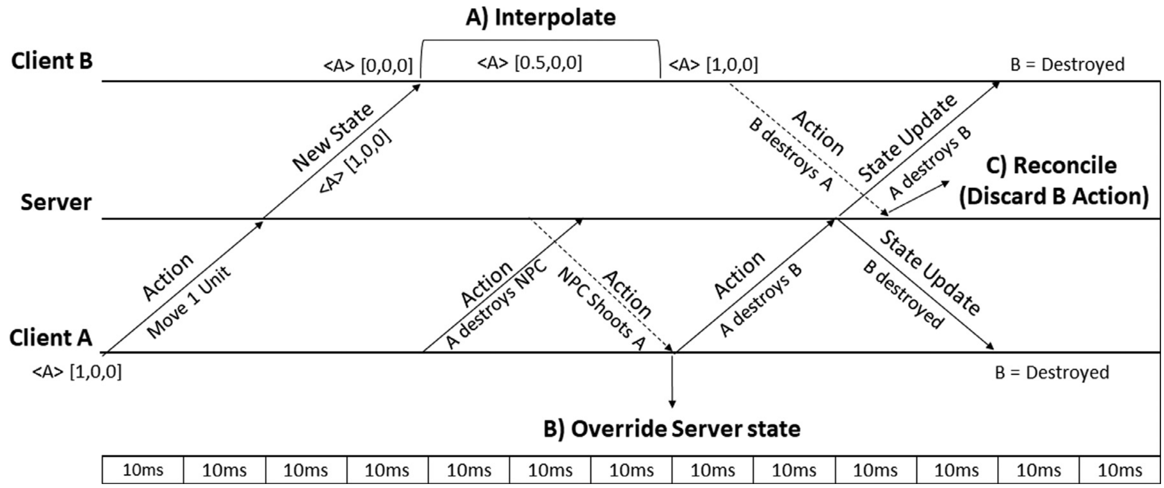


**Fig. 6.** A network action sequence showing techniques used to share actions between clients [44]. In this scenario, two clients perform actions in 10 millisecond (ms) intervals, and each client experiences 20 ms latency. (A) Client side interpolation. When client B receives an update of client A moving. Rather than instantly updating the state on client B, we interpolate between the current and new state over time. This causes smoothing to occur, hiding jumps in poses. (B) Client side overriding. A destroys a Non User Entity (NUE) the same time the NPE attacks client A on the server. Since A's time stamp is placed before the server action, the server reconciles, with client A overriding the state received by the server. (C) Server side reconciliation. Client A destroys client B (sending the result to the server) and 10 ms later, client B destroys client A. The server collects, and reconciles both actions according to the network time stamps (since A happened before B, the server discards B's action).

accuracy by estimating the screen-space distance between the manually marked center of the infrared LED and the manually marked center of the virtual crosshairs. At 5m both systems are at the maximum accuracy, as they are using the shared local map. As the HoloLenses were moved further apart, we synchronized using the native HoloLens tracker's global map and our system continued to load several local maps at 25 m. We measured the error as the pixel displacement between the infrared LED and the virtual crosshairs in screen space (as the screen space visual consistency is all that is required). We then convert this pixel displacement to meters by comparing the known landmarks (two cement pillars) on either side of the dummy (that is placed 2.5 m between each pillar). The results show that at 25 m, the pose resulting from the HoloLens system begins to drift, causing a visual error of $\sim 1$ m at 50 m and about 1.6 m at 75 m. Conversely, our system maintains an accuracy less than 0.5 m at both 50 and 75 m (Fig. 7).

### 4.5. Limitations

The goal of our evaluation was to compare the quality of the alignment of virtual content in screen space over large distances. The results of our evaluation show that our system visually enhances the accuracy of synchronized poses (and therefore enhances the perceived accuracy of placed shared content) between multiple Microsoft HoloLenses in larger-than-room scale environments. However, it still does not achieve the visual accuracy required for LSHF CAR. This is due to two possibilites: innacuracies in the localization system between HoloLenses, and the accuracy of the ICP alignment that directly affects the accuracy of our system. Nevertheless, our improvements are enough to allow indirect representations of users in AR to hide this imprecision, providing the illusion of high fidelity at large scales. Another limitation is that each HoloLens must be initialized within a known starting location, but this can be easily addressed by using GPS to obtain a rough initial position, then loading candidate maps (HoloLens anchors) near the provided coordinates. Finally, we did not provide a complete analysis of the accuracy against a ground truth, because our focus was the quality of the alignment in screen space as users are unlikely to notice depth errors over large distances.

## 5. HoloRoyale: the first instance of a LSHF CAR experience

In this section, we fit our envisioned LSHF CAR experience and our current implementation into the design space outlined in Section 3. HoloRoyale is the first instance of a LSHF CAR experience where several users work together to defend key locations placed in urban areas against an invasion of virtual robots. Users had to form teams and defend several communication satellites distributed in the environment by destroying robots that attack the satellites in waves. After several waves a boss robot appears. The game ends when players destroy the boss robot or the robots destroy at least one base station. This experience leverages the high fidelity features of our system, including visual occlusions and real-virtual world interactions. This experience is also deployable in larger areas and is specifically designed for distributed interactions. One key limitation when applying the gamespace was that the original design of HoloRoyale had to be modified in order to fit our design space. As such, it is likely that when applying other experiences to this design space, their narrative will also need to be modified. This section details these modifications and the applications of the elements within the design space to create the experience. Then, in order to validate the experience, we demonstrated it at several conferences and describe the observations made during the demonstrations.

### 5.1. Fitting the experience to the design space

By fitting HoloRoyale to the established design space (Fig. 3), we apply the suggested configurations, addressing the challenges unresolved by the platform that we implement our experience on.

*Interaction via remote AR avatars:* Although the implementation presented in Section 4 improves the accuracy of pose synchronization over large distances, this error is still noticeable, and can be further impacted by the network latency. To overcome this limitation, we modified how users interact with AR content. Instead of via a virtual hand held pistol-like controller per original design, we represent the users as remote avatars. Each user has two virtual drones that follow them (Fig. 8a). Users interact with the virtual environment through these virtual drones by firing virtual lasers in the direction the user is facing. These avatars provides several key benefits:

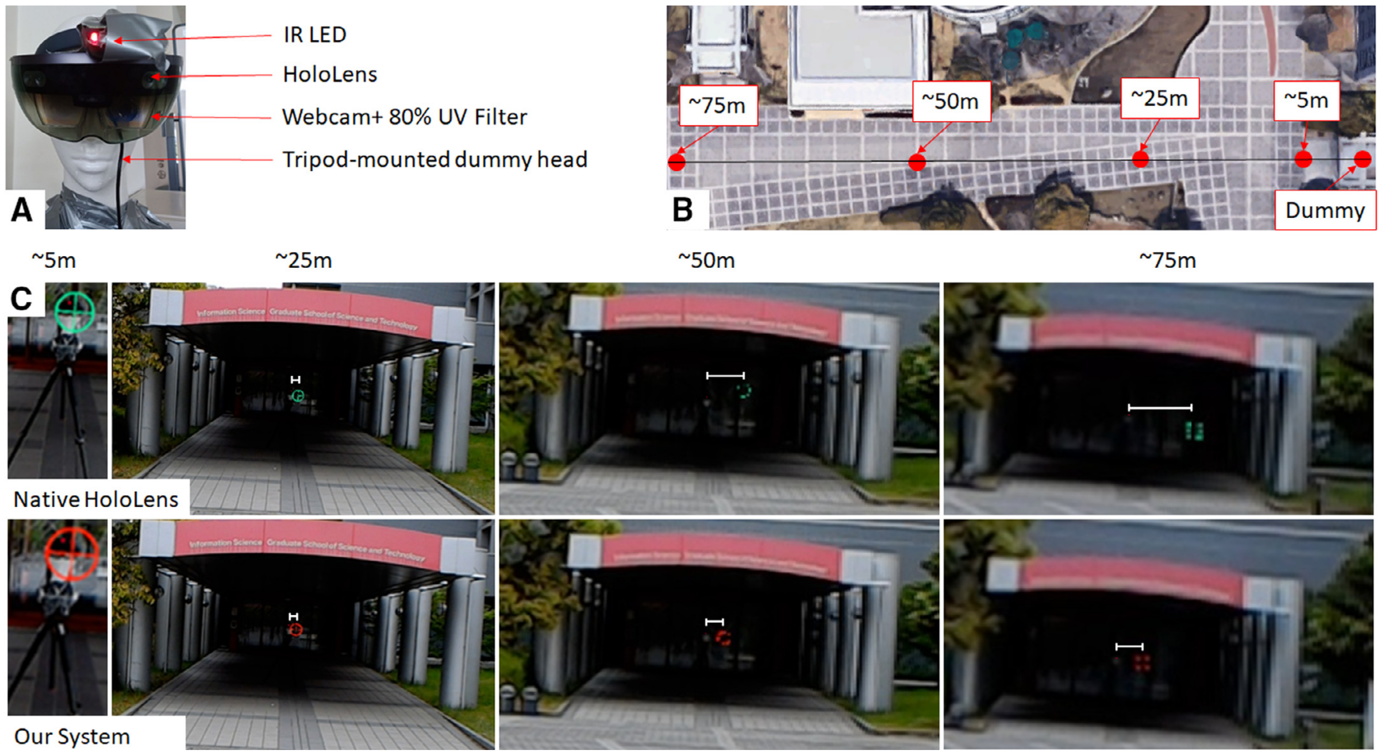


**Fig. 7.** Improved accuracy of synchronized poses between multiple HoloLenses. (A) We mounted two HoloLenses onto dummies, with a webcam embedded into the dummies' eye to take pictures through the HoloLens. (B) We place the dummies in a previously mapped environment and oriented both HoloLenses so that they face each other at distances 5 m, 25 m, 50 m, and 75 m. (C) With increasing distance, the error of the shared position of the out-of-the-box HoloLens system becomes very large, while our system maintains a higher accuracy.

- Hide any inaccurate pose synchronization while still keeping the illusion of perfect tracking between users.
- Hide temporal inconsistencies by utilizing client side interpolation (Fig. 6a) & overriding (Fig. 6b) [44].
- Provide targeting assistance for users and consistent interactions by tagging the target for interaction, and orienting the virtual avatars towards the target of interaction on all clients (Fig. 8c).

Additionally server side reconciliation [44] allows us to resolve conflicting states between users (Fig. 6c).

*Spatial understanding:* We provide a minimal interface (Fig. 9a) to assist with spatial understanding. We place 2D symbolic attractors in the upper compass bar to highlight key gameplay objective locations. These serve two purposes, the first as a directional awareness aid, the second to provide additional information of the game context such as, the distance to the location and the direction relative to the user. We also provide several variations of the WIM [61] that can be zoomed by holding down one of the buttons on the gamepad.

*Communication:* To facilitate communicate between non co-located users, we provide a voice and emoticon communication system. Users can select an instant message by holding one of the buttons of the gamepad, and using the thumbstick to select one of the available messages, then releasing the button to send it. Users can also use voice chat by holding another button and speaking. The UI shows any instant messages sent, and which users are utilizing the voice chat system (Fig. 9a).

*User Redirection:* To navigate users towards key locations, and away from dangerous areas, we provide both 2D symbolic elements in the UI and 3D diegetic user redirection elements. The 2D symbolic elements in the HUD's compass bar flash to remind users

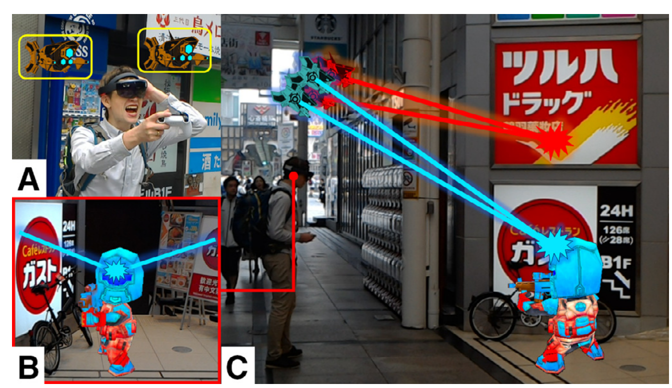


**Fig. 8.** Unattached virtual avatars (A) provide a means of hiding spatial and temporal errors in pose synchronization. (B) If a user interacts with a target (C) by tagging the target for interaction and orienting the virtual avatars towards the target, other users observe a correct interaction (blue) instead of a miss due to spatio-temporal inconsistencies (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of their objective, attracting their attention and guiding them towards their target. A 3D radio portal functions as a diegetic attractor, highlighting where users should be standing. The 3D symbolic navigational cues highlight a suggested pathway towards a target, functioning as both a navigation assistance tool and as a user redirection element (by guiding users towards key locations while avoiding areas tagged as dangerous, or likely to cause our system to fail). The 3D diegetic repellers are synonymous to roadwork barriers, blocking pathways to areas we don't want users to be in (Fig. 10b).


**Fig. 9.** Our user interface and experience provides the user (A) Spatial understanding of their environment (yellow rounded squares), information related to the communication between users (red rounded squares), and elements for navigation and user redirection (green squares). The compass bar at the top of the user's view shows the relative rotational difference from the user's view angle. The arrows highlight the suggested pathways for users, while the roadwork signs indicate an impassable area. The map tool at the bottom provides limited spatial information. The map has three variants: (B) World In Miniature [61], (C) Simplified world in miniature, (D) Radar showing only relative positional information. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

### *5.2. Demonstrations*

To obtain some early observations and feedback, we deployed our experience at UIST and ISMAR [74,75], observing how users played the game. The demonstration area within these conferences was indoors, less than 250 $m^2$ in size, not isolated from pedestrian foot traffic, and featured occlusions from both other demonstrations and the surrounding environment. Although it would have been desirable, we did not collect demographic information of users during the demonstrations. However, since the conferences were in the fields of AR and VR, it is safe to assume users who played HoloRoyale at these conferences were familiar with AR technologies, more so with the Microsoft HoloLens. Each group of 3 users would play the game for 2 min, then return and if desired, provide feedback. The feedback given suggested the game was both fun and the interactions were natural. We noticed that users instinctively respected the user redirection elements with little instruction about them. Even so, some users reported being frustrated at the placement of the repellers, this is likely because of the small area of movement was being restricted further. There were limitations in the venues; The play areas were small ($<$250 $m^2$), significantly crowded, did not distribute users over the play area and therefore was not highlighting the collaboration of the large scale. The sessions were also restricted to 2 min. Because of these limitations the demonstrations did not target our envisioned scenario. The feedback from the participants and the limitations of the demonstration venues raised the question on the effectiveness of repellers and their effects on the user's enjoyment of the experience in LS environments. As such, we conducted a controlled user study, eliminating all possible limitations. We describe the study in the next section.

## 6. Evaluating the navigation effect of diegetic repellers

We expected that a user's instinctive reaction to a virtual diegetic repeller will be analogous to a real wall, inciting them to find an alternate path to their goal. We also expected the virtual diegetic repellers to have no significant impact on the user's enjoyment because users would view the diegetic elements as part of the game experience[64]. During our demonstrations, users obeyed the boundaries created by the virtual diegetic repellers but reported frustration due to the restrictions they created. We hypothesized that this was due to the limited demonstration area. As we conceptualized repellers as a means of user redirection in LS environments, we conducted a user study focusing on the effect of virtual diegetic repellers on user navigation in a LSHF CAR context.

We deployed a variation of HoloRoyale in a 15,625 $m^2$ area on our university campus (see Fig. 10c) and recruited participants to play it in groups of 3 members at a time. For this user study, we removed navigation cues, and restricted spatial understanding tools to the compass bar (for showing attractors, Fig. 9a) and the radar representation of the environment (Fig. 9d). We also slightly modified the system that HoloRoyale is built on, increasing the

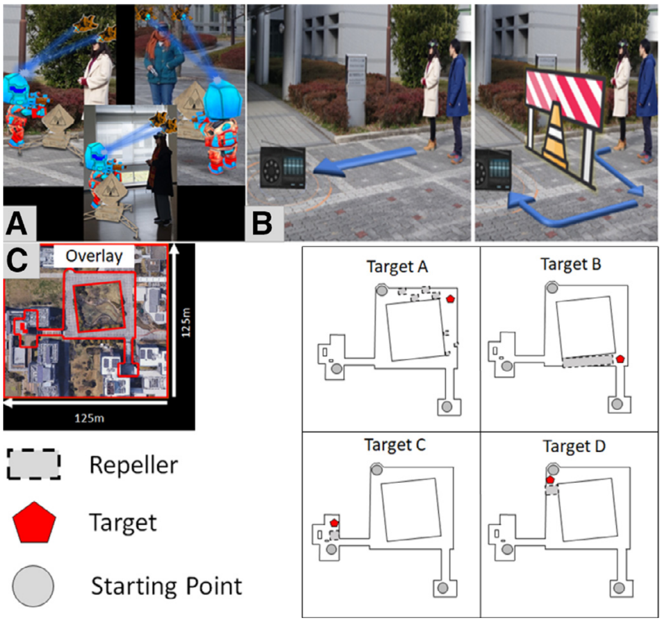


**Fig. 10.** We deployed HoloRoyale on a university campus, spanning a 125 × 125 m area, containing both indoor and outdoor areas. (A) The first part of our user study has participants move to three statically placed bases, defending them against virtual robots. (B) participants proceed to move to one of four target locations, the variable we introduce appears during this phase. This diegetic repeller is represented as a roadwork construction sign. (C) The layout used in the study, we show a layout for each target location as different repeller layouts are setup to simulate different scenarios. [Target A] Virtual barriers in open spaces creating an obstical course. [Target B] A long hallway being barricaded off. [Target C] Virtual navigation in narrow areas. [Target D] Repellers not seen until the last approaching second to attempt a frustrated response. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

client pose synchronization rate to 60 Hz. We represented the virtual diegetic repellers as a construction roadwork sign (See Fig. 1b) and the diegetic attractors as a highlighted radio box. We limit their visibility to 8 and 5 m, respectively. The 2D symbolic in-HUD attractors were visible at all times. We had the following hypotheses:

**H1** Participants will respect the barriers formed by diegetic repellers.

**H2** The virtual diegetic repellers will not significantly impact the participant's enjoyment.

### 6.1. Participants

We recruited 24 participants (17 male, 7 female) between 22 and 34 years (mean 25.5, standard deviation 3.8), from students within our university via email, poster, flyer, and social networking. Participants selected their preferred time slots, creating 8 groups of 3 participants each from overlapping time preferences. Among them, 17 participants had not used a HoloLens before, 8 participants had not played a location based game before, and 15 participants rated their ability to use a map tool as above average.

### 6.2. Procedure

Our study consisted of two phases, a preparation phase and the study itself. We show a timeline in Fig. 11.

Upon their arrival, participants listened to a brief explanation of the experiment procedure, signed consent forms, and filled in a pre-study questionnaire. The participants then took part in an interactive tutorial of HoloRoyale that explained the gameplay, tasks, and user functions (5 mins). The tutorial had several paused sections, allowing participants to familiarize themselves with all game functions.

| 2 Hours | | | | | |
|---|---|---|---|---|---|
| 5min | 45min | 5min | 15min | 45min | 5min |
| Tutorial | Trial #1 | Q&A | Rest | Trial #2 | Q&A |

**Fig. 11.** Experiment timeline.

We assigned each participant one of three bases to defend (see Fig. 10c). Once all participants arrived at their assigned base the game was started. Participants played two sessions of HoloRoyale with the following flow (45 min. each):

(1) *Preparation Phase (Defend):* Participants defend bases by shooting virtual robots. The phase is completed once thirty robots are destroyed at each base. This phase ensured participants were at their respective starting locations before starting the next phase.
(2) *Trial Phase (Upload at target location):* One of four statically placed target points appears in a random order within the play area. Participants converge to the location of the target, standing within 2 m of it. Once all participants arrive at the target point, a progress timer starts to count down, with the phase ending after 10 s.
(3) Participants return to their assigned bases and repeat phases 1 and 2 for all four target locations.
(4) *Final phase* After all locations were visited, the final boss appears at a static location. Participants converge to the boss' location and destroy the boss, ending the session.

After each session, everyone returned to fill in a post-session variant of the Usability Metric for User Experience [76] (See Fig. 13) (4.5 mins).

Between the two sessions, participants took a 15 min rest. After both sessions, participants were free to provide free-form feedback. The total time for each group was approximately 2 h.

For safety reasons, during the user study each participant was shadowed by an assistant. The assistant did not interact with the participant, unless the participant reported something wrong with the system during play (for example, a system failure). This happened during 12 trials and the data for those trials was discarded. We compensated each participant for their time ($\sim$10 USD per hour). This study was approved by the institutional review board of (Removed for Anonymity)

### 6.3. Variables

Our experiment was a *within-subjects* user study with the following independent variables:

**Repellers** $\in$ { Displayed, Hidden }

This describes if diegetic repellers were present in the session. We counterbalanced the order this variable was chosen.

**Target** $\in$ { A, B, C, D }

Each session had four trials, one for each target. Repeller layouts were unique for each target creating the following situations: Barriers in open spaces, a long hallway barricaded off, virtual navigation in narrow areas, repellers that are not visible until a participant is near the goal forcing a long redirect (See Fig. 10). The order the target locations appeared in was randomized and counterbalanced between groups.

**SessionNumber** $\in$ { 1, 2 }

We include the session number to observe if there was a learning effect between sessions.

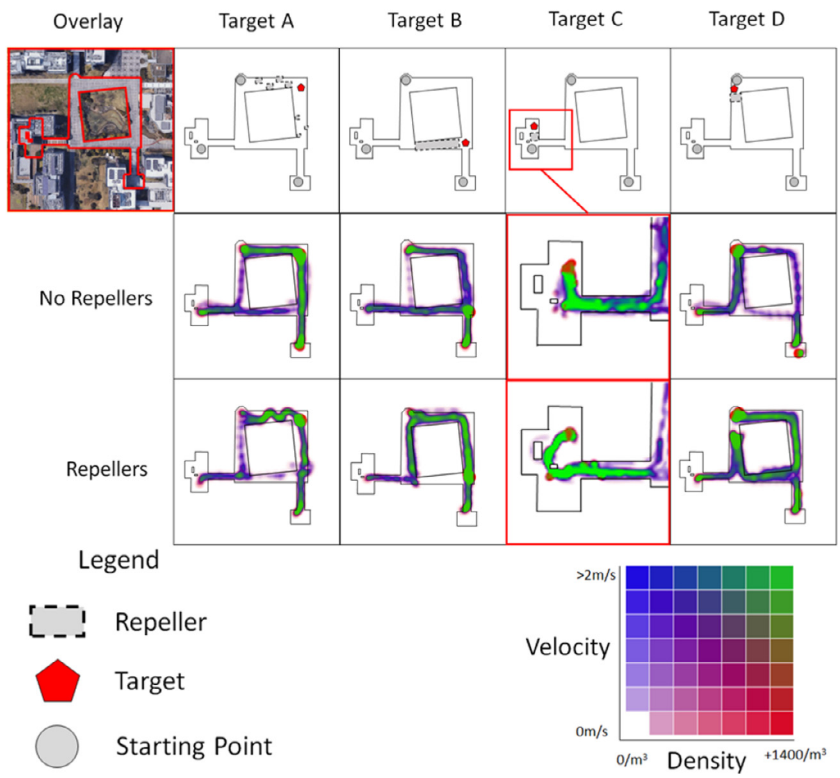


**Fig. 12.** KDE heatmaps showing the density, and velocity of poses we collected during the trial phase of our study. Participants move from their starting position, to one of the target locations that appear in a random order. We use a kernel of 3 m (with the exception of target C, for which we use a 1 meter kernel).

### 6.4. Results

**H1** stated that participants will respect the boundaries set by the diegetic repellers. To investigate this we plot the pose and velocity data recorded during each trial (Fig. 12). We estimate the total amount of poses at ~1,382,400 (average time per target @ 4 mins * 4 targets * 30 poses per second * 2 repeller conditions * 3 participants * 8 groups). We plot the poses as a KDE heatmap with a kernel size of 3 meters for each target, except C that we use a 1 m kernel size, due to the smaller viewport. They show that when repellers are present, participants walked through the barricated areas in 4/216 cases.

**H2** stated that the existence of virtual diegetic repellers will not impact the participants' enjoyment. We investigated this by analyzing the results from the likert questionnaire participants answered after each session, as well as the amount of time participants took to complete the game. We use the criterion of $p < 0.05$ to determine statistical significance.

We show the results of our likert questionnaire in Fig. 13. We compare the answers to our questionnaire with a Wilcoxon Signed-Rank test. The results showed that the presence of repellers had no significant impact on the frustration ($T = 39.5$, $p = 0.394$), ease of use ($T = 17.5$, $p = 0.94$), and how much participants enjoyed the game ($T = 39.0$, p = 0.46). On the other hand, participants reported that repellers significantly affected their ability to perform their intended actions ($T = 21.0$, $p = 0.0096$). The presence of repellers also negatively affected the participants' mental image of their surroundings ($T = 19.5$, $p = 0.046$) and their ability to communicate with their partners ($T = 5.0$, $p = 0.008$).

To investigate if participants reached their targets faster as they became more familiar with the user interface and the game layout we compare the time it took them to finish each session. We show the time participants took to complete each session in Fig. 15. As the Shapiro–Wilk test showed that the data was not normally distributed we used the Wilcoxon Signed-Rank test. The results show that participants completed the second session significantly faster ($T = 47$, $p < 0.001$). We checked how long participants spent looking at the zoomed map tool between sessions. The Wilcoxon Signed-Rank shows that during the second session participants spent significantly less time looking at the map ($T = 96$, $p = 0.005$). Finally, a Wilcoxon signed-Rank test showed that $t_w$, the time taken between sessions minus the time looked at the zoomed map tool, was significantly reduced ($T = 109$, $p = 0.019$).

We also investigated how the presence of repellers affected the time needed to reach each target location. As expected, participants took longer to reach the target when repellers were present (Fig. 14). A Wilcoxon Signed-Rank test showed a significant difference in the amount of time taken for targets A, C, and D ($T = 0.0$, $p = 0.001172$) and no significance on target B ($T = 0.0$, $p = 0.093$).

### 6.5. Discussion

The results of the KDE plot visually *support* **H1**. When repellers were present in the scene, participants mostly respected the boundaries they set. This was the case even when users had to follow a complex pathway in wide areas (Target A), or a maze in a smaller area (Target C). It is also worth noting that the repellers were not 100% successful. In the trials where repellers were not successful, one or two team members who had already arrived at the target location continuously prompted the remaining participants to hurry. One participant even suggested to ignore the repellers and to walk through them, when his teammate could not immediately find the alternate path around the repeller. This suggests that although virtual diegetic repellers present an intuitive barrier that is mostly respected, users may disregard them, e.g., due to peer and time pressure, frustration, or carelessness. When designing LSHF CAR experiences it is thus important to include reinforcing effects that prevent users from walking through diegetic

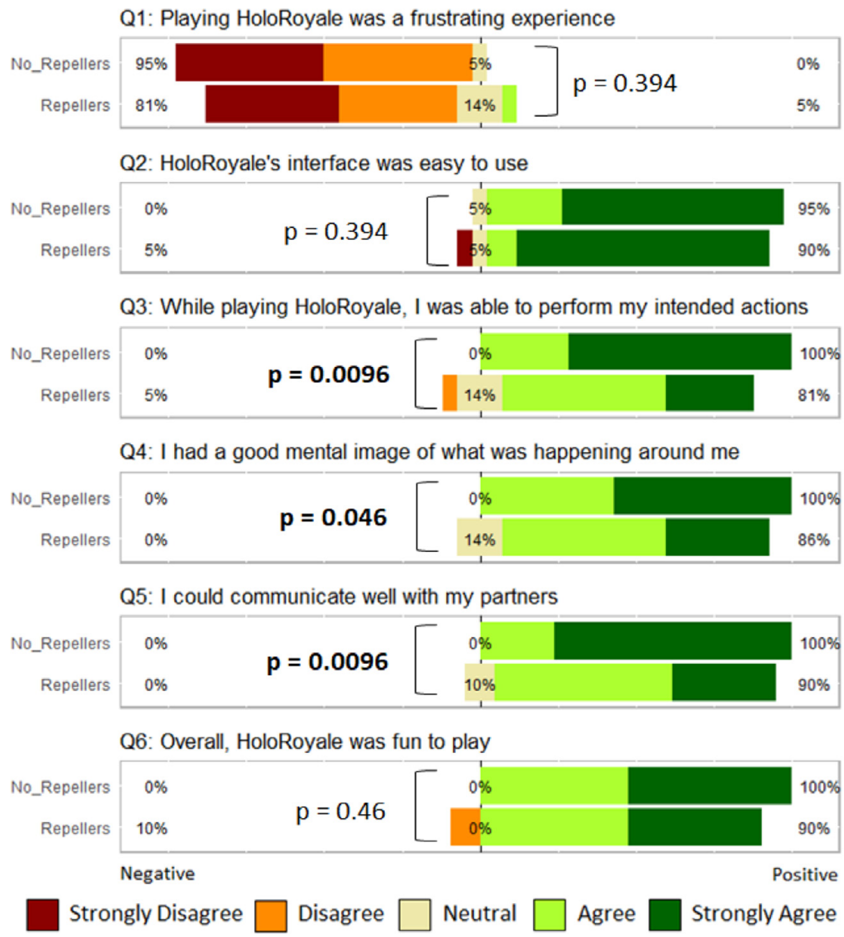


**Fig. 13.** The results from our variation of the usability metric for user experience [76].

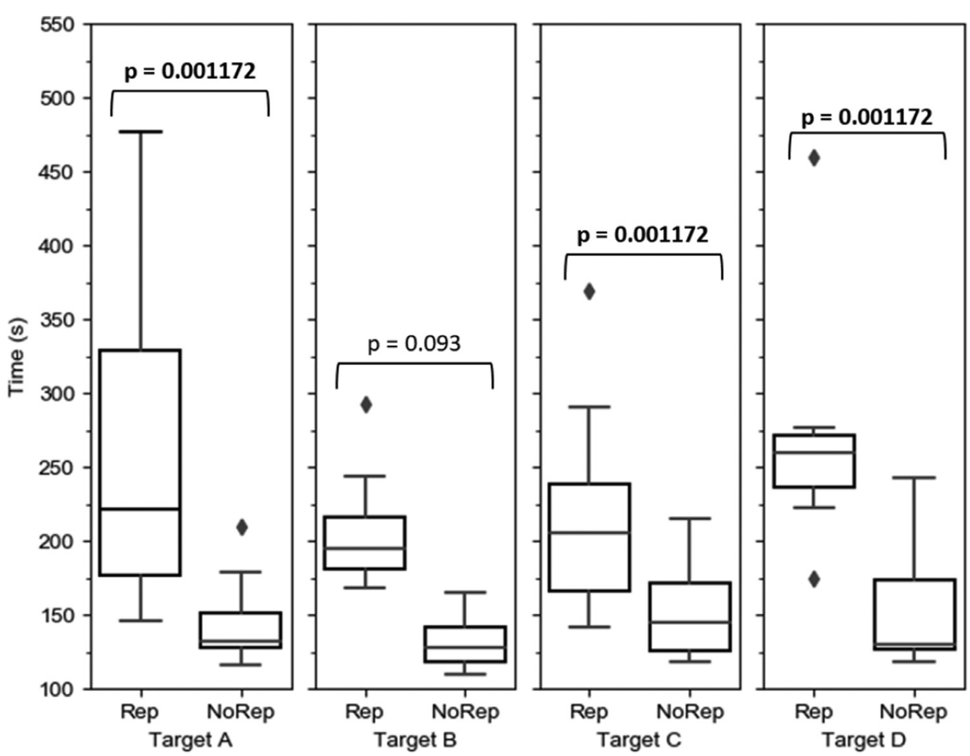


**Fig. 14.** Box plots showing the time differences for each target with/without virtual repellers.

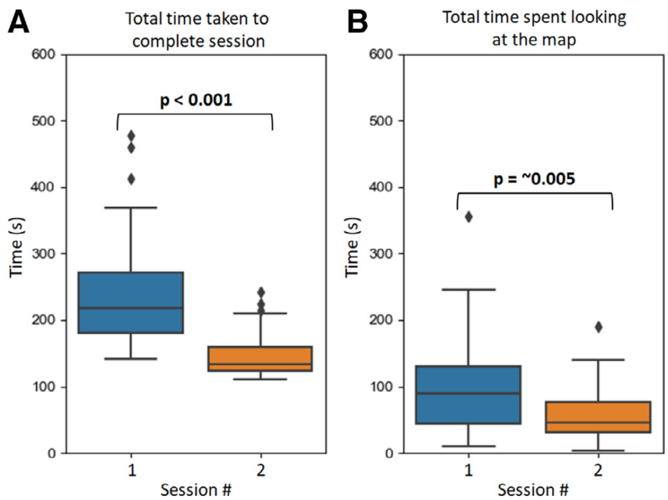


**Fig. 15.** Boxplots showing (A) the time taken between sessions and (B) how long participants spent looking at the zoomed map of their environment between sessions.

repellers, e.g., by turning off the CG and prompting participants to return or by penalizing the crossing of diegetic repellers. Furthermore, when creating LSHF CAR experiences designers need to carefully consider the effects of collaborative mechanics as well as the placements of repellers, attractors, and areas of interest.

The statistical analysis of the likert questionnaire supports **H2**. Although participants did not report a significant impact on their enjoyment of the game or frustration, repellers significantly impacted the answers to questions Q3, Q4 and Q5. It is likely that participants felt that they could not complete the task as intended because the repellers blocked their way and they had to think of an alternative approach. This is supported by the KDE plot for target D, where the repeller in front of the target location forced participants to turn around to find an alternative path. Nevertheless, 81% of the participants stated that they could perform their actions as intended. We found no statistical difference between the answers of users who reported familiarity with the HoloLens vs those for whom this was their first experience, and therefore do not believe this was a factor in the overall results. This also suggests that the designed experience was very natural and could be used by both novices and those familiar with the platform.

The participants' difficulty to create a clear image of the environment when repellers were present could be due to a variety of factors. First, the repellers changed their location for each target. This could have confused the participants and made it more difficult to maintain a clear image of the environment. Second, the repellers were only visible when participants came close to them. This could have also contributed to the participants' anxiety when exploring a path. Furthermore, we did not provide navigation cues that could have helped participants efficiently navigate around the repellers. Nevertheless, 86% of the participants stated that they had a good mental image of the environment.

To our surprise, participants reported a significant negative impact on their ability to communicate with their peers. On follow up interviews several participants stated that it was more difficult to accurately portray and communicate the alternative pathways to reach a target when virtual repellers were present. This suggests that more detailed environment maps, e.g., WIM, that contain information about repellers and navigational cues could simplify communication in complex scenarios.

As expected, participants required significantly less time to complete the second session. This could be in part because participants become familiar with the layout of the environment and the UI. This is supported by 30% of participants with no prior experience reporting that they had initial difficulties understanding how to locate the target areas, but became adept at doing so very quickly. Another explanation of this finding could be the simple layout of our environment, which made it relatively easy for participants to find alternative routes to the target location. The simple layout could have also allowed participants to easily recognize the target location from the indication in the compass bar. These observations are supported by the reduced time participants spent looking at the map as well as the reduction of $t_w$ during the second session.

When providing free-form feedback after both sessions were completed, overall participants stated they enjoyed HoloRoyale and liked having the ability to communicate with each other during the sessions. In addition, several participants with HoloLens experience reported a feeling of a 'larger FOV' when playing HoloRoyale, compared to other applications they have tried previously. This could be because during the preparation phase participants were actively engaged in the game. This focused their attention at the center of the screen thus effectively reducing the noticeability of partially rendered CG due to the limited field of view. At the same time, during the trial phase participants were asked to navigate through a LS environment whilst simultaneously being exposed to UI content being placed along the screen border. As the UI content was visible at all times, this could have reinforced the illusion that the CG was not bound by the HoloLens' field of view. In the future, it is necessary to investigate what prompted this reply from our participants as it could provide means to create immerse experiences on OST-HMD with a limited field of view. Finally, although assistants were not allowed to interact with the participants participants, one of the assistants reported observing that a participant walked into a grassy area outside of the marked play area. This was later determined to be due to an error in the system's tracking during runtime. As a result that participants data was removed from the study.

### 6.6. Limitations

There were several limitations in this study. First, the study focuses primarily on a single representation of a virtual diegetic repeller, and did not evaluate the effects of all design elements adapted from the design space, like the indirect representation for example. It is thus necessary to investigate the effectiveness of other design elements and their effect on navigation and interaction. For example, navigation cues and more detailed maps could help overcome the impact of repellers of the user's mental environment model.

Second, due to the limited battery life of the HoloLens, we could not provide more than four target locations in a session. Furthermore, although we conducted our study in a large environment it was rather simple. As our study was conducted during class time the university campus was also largely without crowds. A more complex environments with many more distractions and pedestrians could lead to different results.

Third, there are technical limitations of the HoloLens display, resulting in transparent rendered content, that could affect the fidelity of the overall experience and should be investigated in future studies.

Fourth, we only investigated one small component of the design elements described in our design space, and therefore should evaluate others, such as the effects of indirect vs. direct interactions in the presence/absence of spatio-temporal inconsistencies in the collaborative environment. Another possible future interaction is the effect of specific methods of communication within the LS interactions.

Fifth, we modified the original gameplay of HoloRoyale, removing any instances of virtual robots during the trial phase to prevent additional factors from affecting the navigational effect of repellers. Additional time pressure to rush to the target location and return to the bases due to the presence of attacking enemy robots, may have prompted participants to ignore repellers more often further underlining the need for reinforcement.

Finally, the environment does not completely match our target scenario. To ensure participants' safety there were no hazards in the area we deployed. It is thus unclear if in an AR context visual diegetic repellers could be sufficient to remind users of the danger thus keeping them out of harms way without the need of reinforcement.

## 7. Conclusion

In this work, we describe a series of requirements for our LSHF CAR experience. From the requirements we established a design space drawing on both technical implementations and design aspects from both AR and video games. By applying the hardware aspect of our design space we created a software architecture and technical implementation that improves the accuracy of synchronized poses between multiple tracking systems. Then we apply our target experience to our established design space, creating HoloRoyale, the first instance of a LSHF CAR experience. Finally, we conducted a user study to explore how virtual diegetic repellers affect user navigation in a LSHF CAR context. The results from the user study suggest that virtual diegetic repellers are effective user redirection elements that do not significantly impact the user's overall immersion.

*Future work*

The work in this paper opens up several new avenues for future work. The first avenue of future work is the application of our established design space into other LSHF CAR experiences, such as outdoor infrastructure planning. To create compelling LSHF CAR experiences over large distances it is also important to investigate the effectiveness of other design elements identified in our design space and their interactions. This includes the effect of indirect interactions on targeting assistance in the presence of temporal-spatial inconsistencies should be investigated. This is especially prominent in LSHF CAR scenarios due to the possibility of interacting with content placed at longer distances (which is highlighted as a problem by [11]). The amount of error users can adapt to when interacting with virtual content before experiencing difficulties is currently unknown and should be explored.

We also plan to investigate the effectiveness of user redirection elements in urban scenarios with a large number of distractors and pedestrians, as well as smaller scale indoor scenes. Our observations also raise questions about the effects the type and density of user redirection elements can have in different scenarios.

Third, to address the participants' comments about the perceived field of view of the HoloLens we need to investigate the effects of UI elements and user immersion on the perceived field of view of an OST-HMD.

Finally, this paper established a crossover between LSHF CAR and video game design spaces. It's possible crossovers between video game design and other AR spaces exist. In the future, we plan to further investigate this crossover, applying the design concepts derived in this paper to other AR domains. One such example is the applications within Virtual/Mixed reality, exploring the thresholds of perceived error under spatio-temporal inconsistencies. Although there is no real-world environment to associate these visual errors with, it's possible that when integrating a multi user shared experience, spatio-temporal inconsistencies can impact user performance. It is possible that indirectly representing users and their interactions can hide these errors, much like in the AR environment. Another example application of our design space could be the use of repellers and attractors to navigate users in virtual reality while avoiding obstacles in the real world.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

This project was supported by JSPS and NRF under the Japan-Korea Basic Scientific Cooperation Program as well as the Creative and International Competitiveness Project Japan. Additionally we would like to thank all participants of the experiment. We also thank Varunyu Fuvattanasilp, Silviya Hasana, Yu Miao, Yuta Uchinimae, and Ao Yanjiao for their assistance during the experiment.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cag.2019.08.007.

## References

[1] Billinghurst M, Kato H. Collaborative augmented reality. Commun ACM 2002;45(7):64–70.
[2] Niantic – pokemon go. 2016. [Accessed 26 June 2017]; https://pokemongolive.com/en/.
[3] Cheok AD, Goh KH, Liu W, Farbiz F, Fong SW, Teo SL, et al. Human Pacman: a mobile, wide-area entertainment system based on physical, social, and ubiquitous computing. Pers Ubiquitous Comput 2004;8(2):71–81.
[4] Kruijff E, Swan II JE, Feiner S. Perceptual issues in augmented reality revisited. In: Proceedings of the IEEE international symposium on mixed and augmented reality; 2010. p. 3–12.
[5] Gruber L, Richter-Trummer T, Schmalstieg D. Real-time photometric registration from arbitrary geometry. In: Proceedings of the IEEE international symposium on mixed and augmented reality. IEEE; 2012. p. 119–28.
[6] Mashita T, Yasuhara H, Plopski A, Kiyokawa K, Takemura H. Parallel lighting and reflectance estimation based on inverse rendering. In: Proceedings of the international conference on artificial reality and telexistence, Tokyo, Japan; 2013. p. 102–7.
[7] Plopski A, Mashita T, Kiyokawa K, Takemura H. Reflectance and light source estimation for indoor AR applications. In: Proceedings of the IEEE virtual reality, Mineapolis, USA; 2014. p. 103–4.
[8] Kán P, Kaufmann H. High-quality reflections, refractions, and caustics in augmented reality and their contribution to visual coherence. In: Proceedings of the IEEE international symposium on mixed and augmented reality. IEEE; 2012. p. 99–108.
[9] Sugano N, Kato H, Tachibana K. The effects of shadow representation of virtual objects in augmented reality. In: Proceedings of the IEEE and ACM international symposium on mixed and augmented reality. IEEE; 2003. p. 76–83.
[10] Rompapas DC, Rovira A, Plopski A, Sandor C, TakafumiTaketomi, Goshiro Y, et al. EyeAR: refocusable augmented reality content through eye measurements. Multimodal Technol Interact 2017;22(4):22:1–22:18.
[11] Piekarski W, Thomas B. ARQuake: the outdoor augmented reality gaming system. Commun ACM 2002;45(1):36–8.
[12] Roboraid – microsoft. 2016. [Accessed 26 June 2017]; https://www.microsoft.com/en-us/p/roboraid/9nblggh5fv3j.
[13] Invaders – magic leap. 2019. [Accessed 22 April 2019]; https://www.magicleap.com/experiences/invaders.
[14] Google – project tango (platform). 2019. https://en.wikipedia.org/wiki/Tango_(platform).
[15] Durrant-Whyte H, Bailey T. Simultaneous localization and mapping: part i. IEEE Robot Autom Mag 2006;13(2):99–110.
[16] Reitmayr G, Schmalstieg D. Mobile collaborative augmented reality. In: Proceedings of the IEEE and ACM international symposium on augmented reality. IEEE; 2001. p. 114–23.
[17] Miyake M, Fukuda T, Yabuki N, Motamedi A, Michikawa T. Outdoor augmented reality using optical see-through HMD system for visualizing building information. In: Proceedings of the 16th international conference on computing in civil and building engineering; 2016. p. 1644–51.
[18] Reitmayr G, Schmalstieg D. Scalable techniques for collaborative outdoor augmented reality. In: Proceedings of the Third IEEE and ACM international symposium on mixed and augmented reality (ISMAR04), Arlington.

[19] Regenbrecht HT, Wagner MT. Interaction in a collaborative augmented reality environment. In: Proceedings of the extended abstracts on human factors in computing systems, CHI'02. ACM; 2002. p. 504–5.
[20] Feiner S, MacIntyre B, Höllerer T, Webster A. A touring machine: prototyping 3D mobile augmented reality systems for exploring the urban environment. Pers Technol 1997;1(4):208–17.
[21] Hanington B, Martin B. Universal methods of design: 100 ways to research complex problems. Develop innovative ideas, and design effective solutions. Rockport Publishers; 2012.
[22] Reitmayr G, Schmalstieg D. Collaborative augmented reality for outdoor navigation and information browsing. NA; 2004b.
[23] Thomas B, Demczuk V, Piekarski W, Hepworth D, Gunther B. A wearable computer system with augmented reality to support terrestrial navigation. In: Proceedings of the second international symposium on wearable computers digest of papers (Cat. No. 98EX215). IEEE; 1998. p. 168–71.
[24] McKenzie G. Gamification and location-based services. In: Proceedings of the workshop on cognitive engineering for mobile GIS; 2011.
[25] Ng G, Shin JG, Plopski A, Sandor C, Saakes D. Situated game level editing in augmented reality. In: Proceedings of the ACM international conference on tangible, embedded and embodied interaction; 2018. p. 409–18. Stockholm, Sweden
[26] Carmack J. Latency mitigation strategies. Twenty Milliseconds 2013. http://www.altdevblogaday.com/2013/02/22/latency-mitigation-strategies/.
[27] Klemm M, Seebacher F, Hoppe H. High accuracy pixel-wise spatial calibration of optical see-through glasses. Comput Graph 2017;64:51–61. doi:10.1016/j.cag.2017.02.001. Cyberworlds 2016; http://www.sciencedirect.com/science/article/pii/S009784931730016X.
[28] Vaughan-Nichols SJ. Augmented reality: no longer a novelty? Computer 2009;42(12):19–22.
[29] Krichenbauer M, Yamamoto G, Taketomi T, Sandor C, Kato H. Towards augmented reality user interfaces in 3D media production. In: Proceedings of the 2014 IEEE international symposium on mixed and augmented reality (ISMAR). IEEE; 2014. p. 23–8.
[30] Sandor C, MacWilliams A, Wagner M, Bauer M, Klinker G. Sheep: the shared environment entertainment pasture. In: Proceedings of the IEEE and ACM international symposium on mixed and augmented reality, ISMAR, 2002; 2002. p. 1–5.
[31] Itamiya T, Tohara H, Nasuda Y. Augmented reality floods and smoke smartphone app disaster scope utilizing real-time occlusion. In: Proceedings of the IEEE international conference of virtual reality (IEEE VR) 2019; 2019.
[32] Weir P, Sandor C, Swoboda M, Nguyen T, Eck U, Reitmayr G, et al. BurnAR: feel the heat. In: Proceedings of the IEEE international symposium on mixed and augmented reality; 2012. p. 331–2.
[33] Sutherland IE. A head-mounted three dimensional display. In: Proceedings of the fall joint computer conference, part I, December 9–11, 1968. ACM; 1968. p. 757–64.
[34] Star wars: Jedi challenges. 2018. [Accessed 22 April 2019]; https://www.lenovo.com/us/en/jedichallenges.
[35] Fujimoto Y, Smith RT, Taketomi T, Yamamoto G, Miyazaki J, Kato H, et al. Geometrically-correct projection-based texture mapping onto a deformable object 2014;20(4):540–9.
[36] Rolland JP, Fuchs H. Optical versus video see-through head-mounted displays. In Fundamentals of wearable computers and augmented reality. CRC Press; 2001. p. 129–72.
[37] Kiyokawa K, Billinghurst M, Campbell B, Woods E. An occlusion-capable optical see-through head mount display for supporting co-located collaboration. In: Proceedings of the IEEE/ACM international symposium on mixed and augmented reality. IEEE Computer Society; 2003. p. 133.
[38] Azuma R, Baillot Y, Behringer R, Feiner S, Julier S, MacIntyre B. Recent advances in augmented reality. IEEE Comput Graph Appl 2001;21(6):34–47.
[39] Wing MG, Eklund A, Kellogg LD. Consumer-grade global positioning system (GPS) accuracy and reliability. J For 2005;103(4):169–73.
[40] Schmalstieg D, Wagner D. Experiences with handheld augmented reality. In: Proceedings of the 2007 sixth IEEE and ACM international symposium on mixed and augmented reality. IEEE; 2007. p. 3–18.
[41] Klein G, Murray D. Parallel tracking and mapping for small AR workspaces. In: Proceedings of the 2007 sixth IEEE and ACM international symposium on mixed and augmented reality. IEEE Computer Society; 2007. p. 1–10.
[42] Engel J, Schöps T, Cremers D. LSD-SLAM: large-scale direct monocular slam. In: Proceedings of the European conference on computer vision. Springer; 2014. p. 834–49.
[43] Mur-Artal R, Montiel JMM, Tardos JD. Orb-slam: a versatile and accurate monocular slam system. IEEE Trans Robot 2015;31(5):1147–63.
[44] Gambetta G.. Fast-paced multiplayer. 2017. [Accessed 27 June 2017]; http://www.gabrielgambetta.com/client-server-game-architecture.html.
[45] Ng K., Ng K.. Wizard's duel ar (prototype): cross-platform multiplayer augmented reality game using google cloud. 2018. https://medium.com/kennethlng/a-wizards-duel-in-augmented-reality-18dbf0cdba7b.
[46] Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement.. J Exp Psychol 1954;47(6):381.
[47] Gutwin C, Vicencio-Moreira R, Mandryk RL. Does helping hurt? Aiming assistance and skill development in a first-person shooter game. In: Proceedings of the 2016 annual symposium on computer-human interaction in play. ACM; 2016. p. 338–49.
[48] Nintendo–Pikmin. 2019. https://en.wikipedia.org/wiki/Pikmin.
[49] 2k studios–xcom: Enemy unknown. 2019. https://xcom.com/.
[50] Microsoft studios–age of empires. 2019. https://en.wikipedia.org/wiki/Age_of_Empires.
[51] Depping AE, Mandryk RL, Li C, Gutwin C, Vicencio-Moreira R. How disclosing skill assistance affects play experience in a multiplayer first-person shooter game. In: Proceedings of the 2016 CHI conference on human factors in computing systems. ACM; 2016. p. 3462–72.
[52] Wang J, Feng Y, Zeng C, Li S. An augmented reality based system for remote collaborative maintenance instruction of complex products. In: Proceedings of the 2014 IEEE international conference on automation science and engineering (CASE). IEEE; 2014. p. 309–14.
[53] Jensen C, Farnham SD, Drucker SM, Kollock P. The effect of communication modality on cooperation in online environments. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM; 2000. p. 470–7.
[54] Klimmt C, Hartmann T. Mediated interpersonal communication in multiplayer video games. Mediat Interpers Commun 2008;309.
[55] Barakonyi I, Fahmy T, Schmalstieg D. Remote collaboration using augmented reality video conferencing. In: Proceedings of the graphics interface 2004. Canadian Human-Computer Communications Society; 2004. p. 89–96.
[56] Bork F, Schnelzer C, Eck U, Navab N. Towards efficient visual guidance in limited field-of-view head-mounted displays. IEEE Trans Vis Comput Graph 2018;24(11):2983–92. doi:10.1109/TVCG.2018.2868584.
[57] 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. ACM; 2008. p. 787–96.
[58] Höllerer T, Feiner S, Hallaway D, Bell B, Lanzagorta M, Brown D, et al. User interface management techniques for collaborative mobile augmented reality. Comput Graph 2001;25(5):799–810.
[59] Stoakley R, Conway MJ, Pausch R. Virtual reality on a WIM: interactive worlds in miniature. In: Proceedings of the CHI, 95; 1995. p. 265–72.
[60] Rompapas DC, Sorokin N, Taketomi T, Yamamoto G, Sandor C, Kato H, et al. Dynamic augmented reality x-ray on google glass. In: Proceedings of the SIGGRAPH Asia 2014 mobile graphics and interactive applications. ACM; 2014. p. 20.
[61] Bell B, Höllerer T, Feiner S. An annotated situation-awareness aid for augmented reality. In: Proceedings of the 15th annual ACM symposium on user interface software and technology, UIST '02. New York, NY, USA: ACM; 2002. p. 213–16. ISBN 1-58113-488-6. doi:10.1145/571985.572017.
[62] Electronic arts–deadspace 3. 2019. https://en.wikipedia.org/wiki/Dead_Space_3.
[63] Bell B, Feiner S, Höllerer T. View management for virtual and augmented reality. In: Proceedings of the 14th annual ACM symposium on user interface software and technology. ACM; 2001. p. 101–10.
[64] Fagerholt E, Lorentzon M. Beyond the HUD-user interfaces for increased player immersion in FPS games. Chalmers University of Technology; 2009.
[65] Holland S, Morse DR, Gedenryd H. Audiogps: spatial audio navigation with a minimal attention interface. Pers Ubiquitous Comput 2002;6(4):253–9.
[66] Wilson J, Walker BN, Lindsay J, Cambias C, Dellaert F. Swan: System for wearable audio navigation. In: Proceedings of the 2007 11th IEEE international symposium on wearable computers. IEEE; 2007. p. 91–8.
[67] Mahmud N, Saha R, Zafar R, Bhuian M, Sarwar S. Vibration and voice operated navigation system for visually impaired person. In: Proceedings of the 2014 international conference on informatics, electronics & vision (ICIEV). IEEE; 2014. p. 1–5.
[68] Adams E. Fundamentals of game design. Pearson Education; 2014.
[69] Lincon PC. Low latency displays for augmented reality; 2017. [Ph.D. thesis].
[70] 6d.ai–cloud based ar. 2018. [Accessed 9 July 2019]; https://www.6d.ai/.
[71] Immersal–cloud based ar. 2018. [Accessed 9 July 2019]; https://immersal.com/.
[72] Nguyen T, Sandor C, Park J. Ptamm-plus: refactoring and extending ptamm. Citeseer; 2010. [Ph.D. thesis].
[73] Sharp GC, Lee SW, Wehe DK. ICP registration using invariant features. IEEE Trans Pattern Anal Mach Intell 2002;24(1):90–102.
[74] Rompapas D, Sandor C, Plopski A, Saakes D, Yun DH, Taketomi T, et al. Holoroyale: a large scale high fidelity augmented reality game. In: Proceedings of the 31st annual ACM symposium on user interface software and technology adjunct. ACM; 2018a. p. 163–5.
[75] Rompapas D, Sandor C, Plopski A, Saakes D, Yun DH, Taketomi T, et al. Holoroyale: a large scale high fidelity augmented reality game. In: Proceedings of the 17th annual international symposium on mixed and augmented reality adjunct. IEEE.
[76] Finstad K. The usability metric for user experience. Interact Comput 2010;22(5):323–7.