2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)

ObserVAR: Visualization System for Observing Virtual Reality Users using Augmented Reality

Santawat Thanyadit [1*] Parinya Punpongsanon [2†] Ting-Chuen Pong [1‡]

[1] The Hong Kong University of Science and Technology
[2] Osaka University

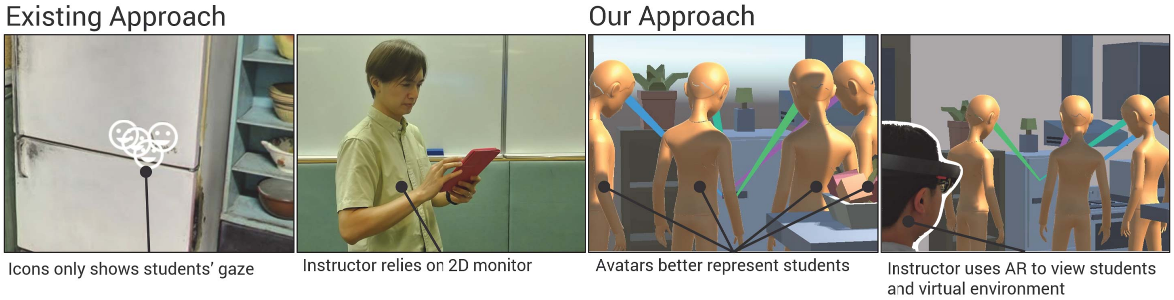

Figure 1: Differences between an existing approach and ObserVAR in observing/guiding students in VEs. **Existing approach** uses an icon to represent each student. The instructor relies on the 2D monitor (image: *Google For Education*). In **our approach**, the avatars represent the students. The instructor uses augmented reality (AR) to observe and instruct each student.

ABSTRACT

While virtual reality (VR) tools provide an immersive learning experience for students, it is difficult for an instructor to observe the students' learning activities in a virtual environment (VE). Thus, it hinders interactions that could occur between the instructor and students, which are usually required in a classroom environment to understand how each student learns. Previous work has added virtual awareness cues that can help a small group of students to collaborate in a VE. However, when the number of students increases, such virtual awareness cues can cause visual clutter and confuse the instructor. We propose ObserVAR, a visualization system that allows the instructor to observe students in a VE at scale. ObserVAR uses augmented reality techniques to visualize each student's gaze in a VE and improves the instructor's awareness of the entire class. The visualizations are then optimized to reduce visual clutter in the scene using a force-directed graph drawing algorithm. In designing ObserVAR, we first investigated visualizations that can provide the instructor with an overall awareness of the VE that can be scaled up as the number of users increases. Second, we optimized the visualization of students by leveraging a graph drawing algorithm to reduce the visual clutter in the class scene. We compared the performance of our prototype with some commercially available user interfaces for VE classrooms. In our study, ObserVAR has demonstrated improvement and flexibility in several application scenarios.

Keywords: Visualization; Remote Collaboration; VR for Education; Asymmetric Interaction;

Index Terms: Human-centered computing—Visualization—Visualization design and evaluation methods; Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed / augmented reality

1 INTRODUCTION

Virtual Reality (VR) technology provides a lot of potential usage in education. For instance, students can attend classes from remote locations and a class instructor can guide students through a VE tour, enabling the students to acquire a rich learning experience through a virtual environment (VE). However, it is difficult for the instructor to observe and instruct the students inside a VE for two reasons. First, the instructor lacks awareness of how to instruct the class because of the environment and the students' activities inside the VE are not visualized for the instructor. The instructor also requires some awareness of his/her surrounding environment for other purposes during the class such as recording students' activities in class. Second, the students can interact with virtual objects without moving closer to the objects in the VE [4, 25]. Thus, the important challenge is to identify the students who are interacting with the virtual objects. In addition, the students' avatars can overlap due to the absence of physical contact, which makes it difficult for the instructor to keep track of the students' activities. As a result, the instructor cannot observe multiple students in the VE.

A traditional method to observe a VR user is by displaying the first-person perspective view through a monitor screen. However, this method does not provide awareness of the VE outside of the VR user's view. Recent research introduced methods that allow an instructor to observe and guide the VR user [12, 17, 19]. A few studies have provided methods for instructors to guide a larger number of VR users (called *one-to-many* guiding tasks). To use this guiding task, the instructor requires an overview of the entire class to observe the students' activities and determine which students requires attention [32]. While Google Expedition [11], a commercial product, allows an instructor to guide students in the VE through mobile devices, viewing the VE through a 2D monitor reduces spatial awareness for the instructor. Moreover, it is difficult for the instructor to associate students with 2D icons (representing students in Google Expedition). If multiple students look at the same objects, which should happen often during a class, the icons can overlap

---

[*] e-mail: sthanyadit@connect.ust.hk
[†] e-mail: parinya@sys.es.osaka-u.ac.jp
[‡] e-mail: tcpong@ust.hk

978-1-7281-4765-9/19/$31.00 ©2019 IEEE
DOI 10.1109/ISMAR.2019.00023

making it difficult for the instructor to distinguish students who are following the class and those who are lagging behind.

This paper proposes ObserVAR, an augmented reality (AR) system that visualizes a VE and students' gaze for the instructor. The students connect to the VE classroom remotely via a network. 3D avatars are used to better represent the students and enhance their social-presence for the instructor. The position of each avatar is optimized to decrease the visual clutter and improve the overall performance. To the best of our knowledge, our work is the first attempt to investigate one-to-many visualization design. In this work, we differentiate the instructor and students using AR and VR systems, respectively, to avoid confusion between the two types of systems. First, we determine the visualization strategies for visualizing the VE and gaze cues that provide the best *overview* and *scalability* for the instructor. Second, we utilize a graph visualization algorithm to reposition the students' avatars in the instructor's environment. The optimization helps to avoid confusion caused by a different VE interaction and improve the instructor's observing experience. Figure 1 shows the difference between the existing approach and our proposed method.

In summary, we provide the following contributions.

- We introduce an AR visualization technique that can scale up with the number of students in the classroom, which improves the interactions between an instructor and a group of students in a VE classroom.
- Through our user study, we determine classroom scenarios in which ObserVAR is suitable.
- We alleviate the occlusion problem, which usually occurs in the scale of a classroom, using a graph drawing algorithm.

2 RELATED WORK

Our work is related to multi-device collaboration techniques that use a combination of devices for shared mixed reality [5, 35]. Such techniques usually require visualization design knowledge that is drawn from the research on visual cues for AR. Other potential techniques with specific devices such as Head-Mounted Displays (HMD) or projection-based AR is beyond the scope of this paper. Please refer to [14, 20] for a review.

2.1 Mixed-device Collaboration

Mixed-device collaboration focuses on creating a shared VE for users with different interaction objectives, using different tools suitable for the users. Previous works have studied the use of mixed-devices in a collaborative guiding task where an AR user guides a VR user using different interactive devices such as tabletop surfaces, HMDs, and handheld ARs [12, 17]. Since only a 2D avatar of the VR user was displayed without indicating the 3D gaze, the AR user had to guess what the VR user could see from observing the avatar. *ShareVR* [15] shows collaborative interactions between a user in a VE (i.e., with HMD) and a user outside such an environment interacting in the same space. While the non-HMD user has a tablet to control and a projector to visualize content from the VE, the gaze and the view of the user in the VE are not visualized, and therefore, the non-HMD user has to infer the HMD user's intentions from his/her physical body. *Mini-Me* [28] presents a visualization technique for enhancing interactivity between AR and VR users in a remote collaboration. A small avatar of the VR user is visualized for the AR user to interact with the VR user in a pointing task. This approach reduces the verbal clarifications and allows two users in different locations to collaborate with each other. Our work extends such a scenario by allowing one AR user to observe multiple VR users (i.e., one-to-many interactions). In such a scenario, the VR users' activities have to be visualized in an effective manner without being occluded by other VR users. Therefore, it is important to investigate any suitable visualization settings to observe multiple VR users (a group of students in this case).

2.2 Visual Cues for Guiding Tasks in AR

One of the main advantages of AR is that visual information can be delivered on top of the physical environment, which allows the users to easily associate relevant information with physical objects. Thus, AR can be used to improve the interactions in navigation, instruction and guiding tasks. In these tasks, the visual cues are introduced to guide users without obstructing the physical environment, and it can be presented according to their purpose.

First, visual cues can be used to guide the attention of users to the objects. Previous works commonly used a line, arrow, highlight, or compass as the visual cue [22, 27, 30]. They used such visual cues to direct a user's attention to a specific object [7]. In our work, we investigate how such visual cues can allow an instructor to observe multiple students in a VE. In particular, we search for a technique that allows the instructor to monitor the activities of the students participating in a VE classroom.

Second, the visual cue can be used to provide information about the objects being viewed by the user. We use this visual cue to lessen the burden of the instructor by displaying only the essential information related to what the students are viewing in the VE. A traditional way to observe the VE is through the first-person view [36], which is commonly used when visualizing VE games. World In Miniature (WIM) [33] allows users to see an entire environment at a glance by viewing it in its miniature version, and has been widely applied in navigation tasks [2, 12] or interaction tasks [26]. Another possible visualization approach is mapping virtual objects into the physical environment, thus allowing an AR user to observe virtual objects in the physical world. This visualization has been used in different applications such as restoring historical sites on top of ruins [24] or seeing through walls [1, 9]. While these visualization settings have been used in different contexts, their effectiveness in the context of multiple VR users has not yet been investigated. In this paper, we investigate the suitability of these visualization settings in the scenario with multiple VR users in a classroom.

2.3 Visual Overload

Aside from the hardware configuration and the visual representations, visual overload also plays an important factor for a smooth immersive interaction. Recent work has been proposed to avoid visual overload by providing only the visual cues relevant to the objects that are of interest to the users working collaboratively in a VE [34]. Other works also suggest that only relevant information should be presented so that it does not overload users with irrelevant information [13, 18]. In our work, we aim to visualize the activities of a group of VR users in a VE where the number of visual cues increases according to the number of VR users. We follow the above-mentioned approaches by using a minimal representation based on a combination of gaze direction and visual display to avoid visual overload. We then perform a user study to find the visualization setting that is scalable according to the number of VR users.

2.4 One-to-many Interactions

Prior research has used visualizations for different types of one-to-one interactions, such as, observing [36], guiding [12, 18], and collaborating [27, 28]. However, visualization that allows an instructor to gain an overview of a group of VR users is still missing from the literature. Such a visualization is required for one-to-many interactions, so the instructor can guide multiple VR users and recognize the VR users' activities at a glance instead of observing each VR user individually. Furthermore, such a visualization would also allow the instructor to recognize the VR users that require the instructor's attention. Therefore, we investigate possible visualizations that are suitable for one-to-many interactions. Our work is the first attempt to investigate the visualizations for one-to-many interactions.

3 VISUALIZATION

In this section, we describe the potential visualizations that allow a class instructor using AR to observe and guide the students in a VE. We start by categorizing visual cues that are important for observing and guiding students in a VE including; (1) gaze and (2) environment.

The *gaze* visualizes where the student is looking in the VE. We selected two gaze visualizations from the instructor's viewpoint (the observer). First, we consider the third person's point of view, which allows the instructor to observe the students from the outside looking in. In this case, geometries such as lines or arrows are used to represent the direction of the students' gaze. Second, we consider a student viewpoint (the VR user), which can usually be seen in a first-person shooter game such as in a 2D video feed from the player. Such visualizations allow the instructor to see exactly what the students are looking at.

The visualization of the *environment* allows the instructor to understand the scene without prior knowledge of the VE. However, it is difficult to visualize the entire VE without occluding the view of the instructor because the instructor might interact with the physical environment. In this paper, we are seeking visualizations that improve the instructor's awareness of the VE. We define three visualization settings for observing the VR users: (1) first person view (FPV), (2) world in miniature (WIM), and (3) world scale (WS). In previous work, those visualization settings are designed for one-to-one interactions. In this work, we adjust each visualization setting so that it is better suited for the visualization of multiple users.

First Person View (FPV)

This method is derived from the first-person shooter game. We visualize what the student is looking at by using 2D video feed directly from the student's headset. The visualization is positioned on top of the avatar's head to indicate which visualization belongs to which student (Figure 2(b)). The VE is not visualized directly, but the instructor can see the VE from the students view. We generated the FPV by copying the students' VE to the PC then created a virtual camera by duplicating the parameter from the VR headset. The gaze direction from the VR-HMD was used to rotate the associated virtual camera. Finally, we rendered the image from the virtual camera as a 2D image on top of the students' avatar head.

World In Miniature (WIM)

Inspired by Stoakley et al. [33], we defined the WIM as the visualization setting which places a mini-avatar representing the student in the miniature world. As shown in Figure 2(c), an avatar that corresponds to the student's position in the VE is placed in the miniature world. To allow the instructor to observe where the student is looking, the direction of the student's gaze is drawn from the head of the avatar to the target position using a green line. The virtual objects the students are gazing at are highlighted in the same color.

We created a miniature version of the VE on top of the heads of students' avatars by scaling down the VE to a physical size of L37.5cm x W37.5 cm x H37.5cm, which was determined from the pilot experiment. To make sure that the mini-avatar was visible to the instructor, the virtual objects between the instructor's line of sight and the mini-avatar were set to be semi-transparent. We determined a line of sight from the instructor's head to the mini-avatar based on the position of the instructor's headset.

World Scale (WS)

This setting directly maps the VE to the real environment. The gaze is visualized based on the location of the student avatars, and the arrow is drawn from the top of the head to where a student is looking (Figure 2(d)). Since the position of the avatars is not aligned with the virtual location, we transformed the gaze direction for alignment to ensure the correct visual cues.

To map the entire VE into the physical room size, we utilized the spatial mapping technique from Microsoft HoloLens [23]. The size of the physical room was used to change the scale of the VE and the orientation of the VE was aligned to match the physical room orientation using the floor layout acquired from the spatial mapping. We visualized the gaze direction q_c of the student by calculating each student's gaze position in the VE P_h. Since the VE that we used is different to the instructor's physical environment due to the spatial mapping M_t, we transform P_h with M_t to obtain the student's gaze position in the physical environment $P_t = P_h \cdot M_t$. Finally, we rendered an arrow from the student's avatar head position to the gaze position. The transformation process is summarized in Figure 3.

Implementation

We implemented our prototype system using Unity3D platform with an Optical See-through HMD (OST-HMD) for the instructor (Microsoft HoloLens) and a VR-HMD for the students (Cardboard VR). The system runs through a PC (CPU Intel I7-7700k 4.20Ghz, 16GB of RAM with NVIDIA GPU GTX1070), which communicates with the HoloLens using a *Holographic remoting* application, it then simultaneously communicates to the VR-HMDs using the android network communication protocol. The VR-HMDs send the gaze directions of the students to the PC via TCP/IP connections and displays the visual cues to the HoloLens based on the selected visualization setting.

4 PRELIMINARY STUDY

During a class, an instructor has to acquire an overview of the entire class by observing the activities of students. To observe a specific activity, an instructor will usually move closer to understand the situation, guide the students and then return to observe the entire class again. From our observation, we believe that the observation system for the VE classroom should cover three different situations; 1) overview, 2) detail and 3) guiding situation. Table 1 shows our defined factors for observing the VE classroom.

Table 1: Important factors in a VE observation system

Factor	Description
Overview	Ability to maintain awareness of multiple VR users
Scalability	Ability to maintain awareness relative to the number of VR users
Detail	Allow an AR user to observe VR users closely to determine the VR users' activities with high accuracy
Guiding	Ability to guide VR users' attention to a target location or position in the VE

While *detail* and *guiding* are parts of one-to-one interactions, which have been well explored in prior research, the *overview* and *scalability* remain a challenge in the context of observing multiple VR users (one-to-many interactions). We conduct a preliminary study to find a suitable visualization setting to provide the *overview* and *scalability* in a VE classroom. The result of this preliminary study will be used to determine a visualization setting in the main study.

4.1 Task

To determine which visualization setting is suitable to provide an overview, we investigated whether participants could observe the general gaze directions of a group of students. We assumed that there are multiple points of interest in the VE where the students' attention is most likely to be focused. The participants were asked to observe the students that looked at these locations to determine those locations where the majority of students paid attention to.

We generate a 3D VE, which is divided into a 2D grid of $250cm \times 250cm$ ($6.25cm \times 6.25cm$ in the case of WIM) marked

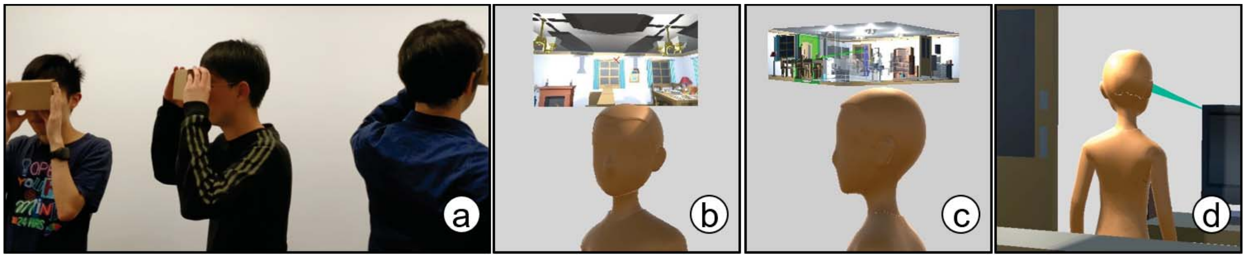

Figure 2: The visualization settings: (a) students in a remote location, (b) First Person View (FPV) visualization, (c) World in Miniature (WIM) visualization, and (d) World Scale (WS) visualization.

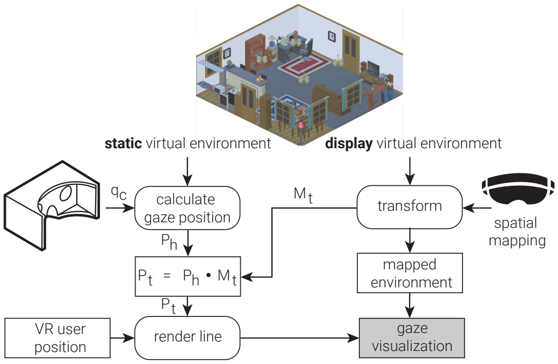

Figure 3: Transformation process of the World Scale (WS) setting.

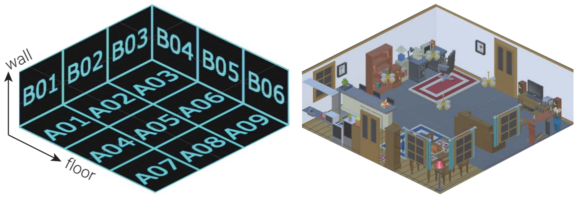

Figure 4: VE in our user studies: (left) the 2D grids marked with letters and number, and (right) a virtual living room

with a letter and numbers (Figure 4(left)) for the participants to determine the grid on which most of the students focused. We performed a study in each visualization setting using groups of 20, 30, and 40 students. To avoid confounding variables and to generate consistent outcomes across visualization settings, we created a data set of students' movement prior to the user study. We recorded movements of volunteers (recruited from the local university) who were exploring the VE according to our instructions. The movement data was then used for the life-size avatars to represent the students in the user study. The avatars were created from Microsoft Paint 3D, and each avatar is $1.6m$ tall, regardless of the physical body of the volunteers in order to remove any external factors from the user study.

4.2 Procedures

Nine participants were invited to take part in the study. The participants performed the task for each visualization setting in groups of 20, 30 and 40 students. We used a with-in subject design for our preliminary study where each participant observed three different sets of recorded data per each group of students; this gave a total of $3 \times 3 \times 3 = 27$ data points per participant. The order of the visualization settings was counterbalanced using the Latin-Square design. We also arranged the data in a balanced manner such that none of the visualization settings benefited from the data. We measured the task completion time and asked the participants to rate the task difficulty with a Single Ease Question (SEQ) [31].

4.3 Results

The results of our preliminary study are shown in Figure. 5.

4.3.1 Accuracy

The overall average accuracy was 78.60%. The WS condition had the best accuracy at 96.30%, followed by the FPV (72.83%), and the WIM (66.67%). A two-way ANOVA showed the main effect of the visualization settings ($F_{2,16} = 6.147,\ p < 0.05$), and the main effect on the number of students ($F_{2,16} = 10.063,\ p < 0.005$). There was a significant difference among the interactions (Conditions x Number of students, $F_{4,32} = 4.536,\ p < 0.01$). Post-hoc analysis was then performed using Tukey HSD and showed significant differences in the WIM settings between 20 and 30 students ($p < 0.05$), and 20 and 40 students ($p < 0.05$). The same post-hoc analysis also showed significant differences between the WIM and FPV for 20 students ($p < 0.05$), WS and FPV for 20 students ($p < 0.05$), WS and WIM for 30 students ($p < 0.05$), and WS and WIM for 40 students ($p < 0.01$).

4.3.2 Time

The mean of the overall completion time was 73.86 seconds. The WS setting was the fastest on average at 43.33 seconds, which was followed by the FPV at 85.8 seconds, and the WIM at 92.19 seconds. A two way ANOVA showed the main effect of the visualization settings ($F_{2,16} = 19.642,\ p < 0.0001$), and the main effect on the number of students ($F_{2,16} = 26.575,\ p < 0.0001$). There is a significant difference in the interactions between the visualization settings and the number of students ($F_{4,32} = 3.773,\ p < 0.05$). Post-hoc analysis performed using Tukey HSD showed the significant differences in the completion time between 20 and 40 students ($p < 0.05$). It also showed the significant differences between WS and FPV ($p < 0.05$), and WS and WIM ($p < 0.05$) for all groups of students.

4.3.3 Single Ease Question (SEQ)

The overall average rating was 3.612 out of 7. The WS setting had the best rating at 5.19 out of 7, followed by the FPV (3.33), and the WIM (2.33). A Wilcoxon signed-rank test showed significant differences between the WS and FPV for 20 students ($Z = -2.047, p < 0.05$), WS and WIM for 20 students ($Z = -2.388, p < 0.05$), FPV and WIM for 30 students ($Z = -2.242, p < 0.05$), WS and FPV for 30 students ($Z = -2.388, p < 0.05$), WS and WIM for 30 students ($Z = -2.692, p < 0.01$), WS and FPV for 40 students ($Z = -2.536, p < 0.05$), and WS and WIM for 40 students ($Z = -2.539, p < 0.05$).

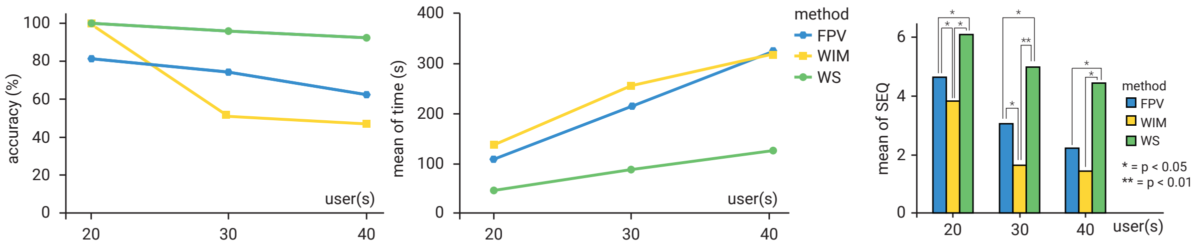

Figure 5: Results of preliminary study: mean of accuracy (left) and mean of completion time (middle), and mean of Single Ease Questionnaire (SEQ) (right).

4.4 Discussion

From the results, we found WS to be the best visualization setting to provide an overview in the VE classroom. The participants spent less time in the WS setting than the other settings, while maintaining more than 92.5% accuracy. Moreover, the participants also rated the WS as the easiest setting across all sets of students in this task. We observed that the participants directly observed the VE while ignoring the students in the WS setting, and once they noticed a possible location, they counted the number of students looking at that area and compared them with other locations. Thus, the participants could decide the exact location where the students were paying attention to. On the other hand, the FPV and WIM showed a noticeably lower accuracy. In these visualization settings, the participants had to observe one student at a time because these settings separately visualized students' VE. We observed that the participants repeatedly lost their progress due to the fact that they forgot the locations of where the previous students were looking. Therefore, they spent more time to complete the task, which was then ranked as being harder than the WS setting.

In summary, the results showed that the WS setting can scale with the number of students to provide the best overview of the three visualization settings. In particular, it allows an instructor (an AR user) to observe up to 40 students (VR users) and filter a group of students with an accuracy of more than 92.5%. Thus, it is clear that the WS setting should be used as "overview mode" in the proposed ObserVAR systems. Other modes such as detail and guide, a prior study [36] and our supplementary study[1] show that FPV can be used to provide the students' view for the instructor, providing the best fit when the instructor wants to observe a student's situation in detail. Thus, FPV should be used as a "detail mode". When the instructor wants to provide detailed guidance to an individual student, WIM is suitable for visualizing the VE of a specific student. WIM suitability with a guidance task has been suggested in our supplementary study and previous researches [2, 12]. We summarized the recommended scenarios for each visualization setting in Table 2.

Table 2: Recommended scenarios for each visualization setting

Visualization setting	Recommended scenarios
WS	Provides an overview of the class for the instructor, allows the instructors to recognize groups of students
FPV	Observes a student activity in detail
WIM	Provides a detailed guidance for individual students

5 REDUCE VISUAL CLUTTER

While the WS setting was found to be the best visualization setting for class scalability, adding multiple visual cues into the VE can cause occlusions (*visual clutter*). The students' gaze visualizations could overlap and cause the scene to become complex for the instructor. In addition, the students can also be too close to the virtual objects, blocking the instructor's view. We reduce this problem by optimizing the location of each student's avatar in the VE of the instructor's view. To do so, we adopt aesthetic rules from the *2D graph drawing technique* [3, 29] and previous suggestions made by participants in our preliminary study (Section 4). The graph drawing techniques are used since they share a similar goal, which is to reduce the visual clutter and improve graph drawing aesthetics. We propose the following aesthetic rules:

- The avatars are positioned at a proper distance from the virtual objects and the proper distance is determined based on an available space, other avatars, and the instructor's position. This is comparable to *uniform edge length* in graph drawing's aesthetics rules (the proper distance is calculated using a formula in the following Section 5.2).
- The avatars that look at the same objects are placed near each other, similar to a cluster in graph drawing.
- The avatars are not occluding the view of the instructor to the virtual objects.
- The gaze visualization does not cross other gaze visualizations or other students' avatars. This is derived from the aesthetic rules in the graph drawing that minimize the *edge crossing* [3, 29].
- The avatars are not overlapping each other.
- The avatars face the virtual objects in their direction of gaze.

To optimize the avatars' location according to the above aesthetic rules, we utilize a *force-directed drawing algorithm* [8,10,21], which are known to produce a crossing-free, symmetrical, and aesthetically pleasing layout. This algorithm[2] works in a typical classroom with less than 40 students. In addition, this algorithm assumes a straight-line drawing, matching the students' gaze visualization.

We formulate the optimization equation by treating the students' gazes as *edges* in the graph, avatars (virtual representation of the students) as *nodes*, and virtual objects as *virtual nodes* (as found in [8]). The virtual objects are treated as virtual nodes as we assumed the virtual objects to be static in our VE scenario. If an avatar is looking at a virtual object, an edge is created from the avatar to the virtual object, each edge has a *spring force* f_s that affects the avatar. There are *repulsion forces* f_r between all avatars. In addition, we calculate the *positional force* f_p, which are modeled from the position of the instructor and the virtual objects in the VE.

The total force at an avatar v is

$$F(v) = \sum_{(u,v)\in V\times V} f_r(\delta_{uv}) + \sum_{(u,v)\in E} f_s(\delta_{uv}) + f_p \tag{1}$$

where δ_{uv} denotes a vector from u to v. The total force is divided into three types of force, repulsive force, spring force and positional force. We describe each force in detail below.

---

[1] please see supplementary materials

[2] pseudo-code in the supplementary material

Table 3: Cube locations

Locations	Descriptions
Center	The cube is in the middle of the room, suspended in air at 1.5 m from the ground, to test the objects that have 360 degrees space around them.
Wall	The cube is on the wall, 1.5 m from the ground, to test the objects that have 180 degrees space around them.
Corner	The cube is in the corner of the room, 1.5 m from the ground, to test the objects that have 90 degrees space around them.
Floor	The cube is in the middle of the room, placed on the floor, to test the objects that are below the users' eye level
Ceiling	The cube is in the middle of the room, 3.0 m from the floor, mounted on the ceiling, to test the objects that are above the users' eye level
AR users	The cube is where the AR user position to test a special situation when the students look at the instructor.

5.1 Repulsive force f_r

The repulsive force follows the inverse square law, which can be written as

$$f_r(\delta_{uv}) = \frac{k_r}{\|\delta_{uv}\|^2} \cdot \frac{\delta_{uv}}{\|\delta_{uv}\|} \tag{2}$$

where k_r is the strength of the repulsion force that ensures each avatar does not overlap each other.

5.2 Spring force f_s

The spring force follows Hooke's law, that is, the spring force is the difference between edge distance δ_{uv} and spring length l_{uv}.

$$f_s(\delta_{uv}) = k_s(\|\delta_{uv}\| - l_{uv})\frac{\delta_{uv}}{\|\delta_{uv}\|} \tag{3}$$

Appropriate Spring Length l_{uv}

Unlike in the graph drawing, a spring length directly affects a distance between avatars and virtual objects, which directly impacts the instructor's experience. If the spring length is too short, the avatars could occlude one another and the virtual objects. Conversely, if the spring length is too long, the instructor would require more head movement to observe the scene. Furthermore, the spring length is also affected by the available space around the virtual objects and the number of avatars in the space.

To determine an appropriate spring length between the avatars and the virtual objects under different circumstances, we conduct a short study to formulate a spring length equation. Eight participants (6 males and 2 females, age: $M = 27, S.D. = 3.5$) were invited from the local university. Five participants had prior experience with VR or AR. All participants had normal or corrected-to-normal vision.

We consider four factors that could affect the spring length, which are 1) number of students that are interacting with virtual objects, 2) available space around the virtual objects, 3) the vertical position of the virtual objects and 4) the type of virtual objects. The spring length affects aesthetics of the scene, and the aesthetics preference can vary from one user to another. Therefore, in this study, we aim to find acceptable ranges instead of precise values. We used a $0.25m^3$ cube as a virtual object and changed its position across six locations; center, wall, corner, floor, ceiling, and AR users, respectively (see Table 3 for details).

For each location, the participants were asked to position themselves where they felt comfortable to observe the scene. We then asked the participants "do you feel that the avatars are too close to the object?". The spring length value was decreased by $0.5m$ until the participants felt the avatars were too close or too crowded. Then, we recorded the spring length value. Next, we increased the spring length by $0.5m$ until the participants felt the avatars were too far from the cube. The study tested each location with 5, 15, and 25 avatars in order.

We found that the number of avatars and available space around the virtual object had a direct impact on the spring length. We then formulated the spring length l_{uv} as a function of the number of avatars n and available degrees around the virtual object r in radians.[3]

$$\begin{aligned} l_{uv}(n,r) = 2.261 + 0.2094n - 0.02273r - 0.00137n^2 \\ -0.008359 \cdot n \cdot r + 0.1534r^2 \end{aligned} \tag{4}$$

where $SSE = 0.1196$, $R - Square = 0.9115$ and $RMSE = 0.1997$. We used the above equation to adjust the spring length l_{uv} in Equation 3. If the students are looking at the instructor, the spring length is a function of the number of avatars n.

$$l_{uv}(n) = 0.4456 + 0.2031n - 0.003435n^2 \tag{5}$$

5.3 Positional force f_p

The positional force f_p is a set of forces that moves the avatars out of undesirable positions. It is composed of boundary force f_b, occlusion force f_o, alignment force f_a and collision force f_c.

Boundary force f_b: The boundary force is used to prevent avatars from moving too close to the boundary of the area. We calculate the shortest distance from the avatars to each boundary δ_b, then move the avatars in the opposite direction $\vec{b}$ following the inverse square law.

$$f_b = \frac{k_b}{\|\delta_b\|^2}\frac{\vec{b}}{\|\vec{b}\|} \tag{6}$$

Occlusion force f_o: To prevent the avatars from occluding the virtual objects, we calculate lines from the virtual objects to the instructor's position. Then, we calculate the shortest distance δ_o from these lines to the avatars' position. We move the node in the opposite direction $\vec{o}$ from these lines following the inverse square law.

$$f_o = \frac{k_o}{\|\delta_o\|^2}\frac{\vec{o}}{\|\vec{o}\|} \tag{7}$$

Alignment force f_a: This force aligns the avatars based on the available space around the virtual object. First, we search the space around the virtual object. If collisions happen, we use collision points and normal vectors to determine an available space vector $\vec{s}$, which is a normalized vector points towards the available space center. We calculate two vectors $\vec{b_r}, \vec{b_l}$, which start at the gaze location on the virtual object and point to the available space boundaries. If a collision is not found, a line is drawn from the gaze location to the instructor, and the two vectors $\vec{b_r}, \vec{b_l}$ are set to be perpendicular to this line. We then calculate another vector $\vec{b_n}$, which starts from the gaze location to the avatars. We find a degree θ from vector $\vec{b_n}$ to boundary vectors $\vec{b_r}, \vec{b_l}$, and add force to the avatars using the following equations

$$f_a = k_a 0.9^{\theta}\frac{\vec{s}}{\|\vec{s}\|} \tag{8}$$

With 0.9^{θ}, the avatar that is out of available space ($\theta < 0$) is heavily penalized, while the avatar that is within the available space is not penalized ($\theta > 0$).

Collision force f_c: Avatars can overlap with virtual objects, which may appear unnatural for the instructor. We formulate a force to

---

[3] The average spring length is shown in the supplementary material.

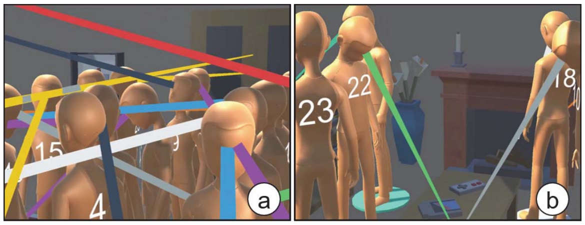

Figure 6: (a) before reducing visual clutter (b) after reducing visual clutter with the same data.

move the overlapping avatars away from the virtual objects. We find the closest boundary of the overlapping objects and calculate a vector $\vec{v}_c$ which starts from the avatar position to the closest point on this boundary. The force is formulated as the following equation.

$$f_c = k_c \frac{\vec{v}_c}{\|\vec{v}_c\|} \tag{9}$$

5.4 Recommended Parameters

We determine a spring length value $l_u v$ through a user study due to the fact that the spring length value has a direct impact on the avatars' position. On the other hand, other coefficients have an indirect impact on the avatars' position, and minor changes in these coefficients do not significantly change the avatar's positions. We determine these values based on our pilot study and observation; k_r, k_s, k_b, k_o, k_a, and k_c, are set to 0.15, 2.0, 1.0, 0.1, 0.5, and 0.2, respectively. We limit the iterations count to 10 iterations per frame to produce approximately 50 frames per second when connected to the HoloLens.

5.5 Avatars Movement

We limit the movement speed of the avatars ($2m/s$) to allow the instructor to track the avatars' movements, when the students' gazes move from one location to another. The movement of the avatars can cause two observation issues. First, it can cross the instructor's field-of-view, which can generate unwanted occlusion. If the path crosses the instructor's field-of-view, but the avatars current location and target location are outside of the instructor's field-of-view, the avatar's path is redirected. Second, the avatar can collide with the instructor, which can create an unpleasant experience for the instructor. We avoid this issue by assuming a circle with the instructor as the center. If the avatars' paths cross this circle, the avatars' paths are redirected to the edge of the circle instead.

6 USER STUDY

The optimized avatars of the students in the previous section are then used in the overview mode of ObserVAR (Figure 6b). Figure 6 shows the comparison between the before and after of reducing visual clutter with our proposed method. We evaluate ObserVAR in the overview mode by considering real world usage cases of VE classrooms.

6.1 Scenario

We assume that instructing in VE classrooms entails a similar process to leading a field trip in a science museum. A science museum is often divided into multiple rooms, with each room having multiple interactive stations, so students can learn by interacting with the different stations. In the VE, each station is comparable to a virtual object. Upon arrival, the instructor introduces to the students the topic related to each room. At the same time, the instructor has to monitor the students' attention. Some students may wander off to observe the exhibits within the room, some students may remain focused and listen to the instructor. After the introduction, the instructor can divide the students into smaller groups and assign them to different interactive stations. The instructor has to know whether the students are at the correct stations. At each station, some students may interact with the station and some students may just observe while waiting for their turns. Thus, the instructor should be able to recognize the active students. Since there are multiple groups of students, the instructor should be able to identify each student with a glance. Once students have finished their interactions, the instructor can ask them to head to a different station. Finally, once students have visited all the stations, the instructor gathers all of the students to conclude the lesson. However, some students may still be lagging behind, therefore the instructor has to identify who they are and call on them.

From this scenario, we identify four important pieces of information for the instructor in the VE classroom.

1. The instructor should be able to locate the positions the students are paying attention.
2. The instructor should be able to track students' attention from one location to another.
3. The instructor should be able to discern whether students are watching or interacting.
4. The instructor should be able to identify those students who have not paid attention to him/her.

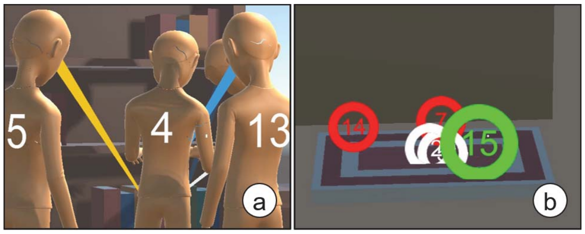

Figure 7: Two conditions in our user study (a) ObserVAR condition, and (b) baseline condition.

6.2 Tasks

We compare ObserVAR (OB) as the proposed method (Figure 7(a)) with the user interface is derived from Google Expedition [11] as the baseline condition (BL) (Figure 7(b)). As we were unable find previous study that tackles visual clutter issues in a one-to-many situation to directly compare with ObserVAR, we chose Google Expedition as the closest and freely available system, which has been used by millions of students [6], for comparison purposes. In the BL, the location of the students' gaze is showed as a labeled circular icon. The size of the icon is set in our pilot study to make sure that it is readable by the participants from a distance. We also highlighted the icon in green and the size of the icon is increased when the participant gazes at the icon. A virtual living room environment (Figure 4(right)) was used in our study. There were 25 students in the scene, each student was represented by an avatar as if each student attends the VE classroom from remote locations. The direction of each student's gaze moved according to a set of prerecorded data, which was generated by recording the movement of a group of the volunteers who were exploring the same VE according to our instructions. The order of the experiment conditions were counterbalanced using the Latin-Square design. We arranged the data in a balanced manner such that the conditions did not benefit from the data. We designed four types of tasks to capture information that is considered important for instructors in VE classrooms – namely, tracking task, observing task, interaction identification task, and attention calling task.

Table 4: Usability rating (OB: ObserVAR, BL: baseline, Task 1: tracking task, Task 2: observing task, Task 3: interaction identification task, Task 4: attention calling task)

#	Details	Cond.	Task 1			Task 2			Task 3			Task 4		
			M	SD	p	M	SD	p	M	SD	p	M	SD	p
Usability- Task Difficulty (SEQ)														
Q1	Overall the task was	OB	5.00	1.60	0.272	6.16	0.58	0.003	6.17	0.72	0.051	6.00	0.74	0.006
		BL	4.58	1.73		3.50	1.45		4.67	1.78		4.67	1.50	
Usability- Subjective Mental Effort Questionnaire (SMEQ)														
Q2	Rate your mental effort	OB	26.67	27.83	0.674	9.16	7.93	0.002	10.83	9.00	0.029	12.08	9.88	0.006
		BL	31.67	25.44		49.17	29.61		30.83	26.36		30.83	25.12	
Usability- Enjoyment and Level of Focus on the Task														
Q3	I enjoyed the experience.	OB	5.33	1.50	0.606	5.83	1.27	0.005	6.17	0.94	0.041	6.17	0.83	0.072
		BL	5.75	1.14		4.75	1.06		5.33	1.30		5.00	1.41	
Q4	I was able to focus on the task activities.	OB	5.33	1.50	0.614	6.50	0.52	0.002	6.50	0.52	0.030	6.33	0.78	0.069
		BL	4.92	1.93		4.50	1.62		5.17	1.70		5.08	1.93	

Tracking task: We asked the participants to track a selected student's gaze by determining the virtual objects that the student was looking at and record all the virtual objects onto an answer sheet. Each student stopped and looked at each object for approximately 10 seconds before moving to the next object, looking at five objects in total and taking approximately one minute to complete the sequence. The maximum speed of avatars in the OB and icons in the BL were set to the same value ($2m/s$). In order to simulate a real scenario, there were also other students exploring the VE simultaneously while each participant tracked only one selected student.

Observing task: After the tracking task, we assumed that the students would start paying attention to specific virtual objects similar to the interactive stations in a science museum. The students' attentions were divided among five virtual objects. We asked the participants to identify all the students' numbers and the virtual objects that the students were looking at to measure the completion time.

Interaction identification task: From the observing task, the participants were asked to identify whether the students were interacting with or just looking at the virtual objects. In the OB, the avatars of the students interacting with the virtual objects moved closer and reached out towards the virtual objects as shown in Figure 7(a). In the BL, the icons of the interacting students changed color from white to red as shown in Figure 7(b). There were randomly two or three students interacting with each virtual object.

Attention calling task: At the start, most of the students turned to look at the participant to simulate scenarios when the instructor calls the students for attention. The participants were asked to identify the remaining students who were not looking at the participant. Five out of 25 students were setup to look at different positions in this task.

6.3 Procedure

Twelve unpaid participants (8 males and 4 females, age: $M = 27.08, SD = 4.54$) were recruited from local universities. None of the participants were familiar with visualization using an OST-HMD device. We performed a user study using a with-in subject experiment design where the participants performed each condition, one after another. The participants were asked to assume they were instructors who were observing a group of students studying in VE classrooms. A short pre-experiment questionnaire was distributed to gather the participants' general information. The participants had two minutes to get accustomed to the VE and the user interfaces. The participants performed each task in sequence, after each task the participant rated the task's difficulty (SEQ [31]), level of enjoyment and level of focus (on a 7-point Likert scale), and Subject Mental Effort Questionnaire (SMEQ [37] from 0:"Not at all hard to do" to 150:"Tremendously hard to do"). Once the participants finished one condition, the participants were given a two minute break before proceeding to another condition. Post-experiment questionnaires were given to the participants to rate social-presence (co-presence, attention allocation, perceived message understanding on a 7-point Likert scale [16]) and their preferences (OB, BL or "no preferences"). Semi-structured interviews were then conducted. The entire experiment took approximately 30-45 minutes.

Table 5: Social-Presence rating (OB: ObserVAR, BL: baseline, 1-Fully disagree∼7-Fully agree)

Social-Presence (Post-Experiment)				
Details	Cond.	Mean	SD	p
Co-Presence				
OB: $M = 6.50, SD = 0.55$, **BL**: $M = 4.58, SD = 1.42$				
I noticed my students	OB	6.58	0.51	0.007
	BL	4.92	1.51	
My students' presence was obvious to me	OB	6.58	0.51	0.003
	BL	4.16	1.64	
My students caught my attention	OB	6.58	0.65	0.003
	BL	4.67	1.16	
Attentions Allocation				
OB: $M = 4.22, SD = 1.73$, **BL**: $M = 3.33, SD = 1.51$				
I was easily distracted from my students when other things were going on	OB	3.42	1.97	0.206
	BL	4.33	1.67	
I remained focused on my students throughout our interaction	OB	5.38	1.24	0.013
	BL	4.58	0.90	
My students did not receive my full attention	OB	3.50	1.08	0.258
	BL	4.25	1.14	
Perceived Message Understanding				
OB: $M = 5.21, SD = 1.79$, **BL**: $M = 3.50, SD = 1.80$				
I understood where my students were focused	OB	6.50	0.67	0.007
	BL	4.92	1.73	
My students' thoughts were clear to me	OB	4.83	2.25	0.016
	BL	3.42	1.51	
It was easy to understand my students	OB	5.50	1.68	0.004
	BL	3.67	1.49	
Understanding my students was difficult	OB	3.00	1.41	0.11
	BL	5.00	1.15	
Preference				
Which condition do you prefer?	OB	100%(12)		0.001
	BL	0%(0)		

6.4 Result

6.4.1 Quantitative Evaluation

Accuracy: Figure 8 shows the accuracy of the participants' performance in each task. A Wilcoxon sign-Rank test showed no significant difference in accuracy of the tracking task ($Z = -0.586, p = 0.558$) and interaction identification task ($Z = -2.070, p = 0.068$). The OB had significantly better accuracy than the BL for the observing task ($Z = -2.673, p < 0.01$) and the attention calling task ($Z = -1.826, p < 0.05$).

Task Completion Time: Figure 8 shows the completion time for each task. A Shapiro-Wilk test showed that our data followed the normal distribution, thus we used analysis of variance (ANOVA) to analyze the task completion time. The participants completed the observing task and the attention calling task in the OB significantly faster than the BL ($F_{1,23} = 22.674, p < 0.01$, $F_{1,23} = 9.501, p < 0.01$). However, no significant difference was found in the interaction identification task ($F_{1,23} = 2.054, p = 0.166$).

6.4.2 Qualitative Evaluation

We used the Wilcoxon signed-rank test to analyze the following questionnaire results.

Task Difficulty (SEQ): In the tracking task, the participants rated the OB and the BL similarly ($Z = -1.098, p = 0.272$). However, in the other tasks, the participants rated the OB to be significantly easier than the BL condition (observing task: $Z = -3.089, p < 0.01$, interaction identification Task : $Z = -2.219, p < 0.05$ and attention calling task: $Z = -2.395, p < 0.01$).

Subjective Mental Effort Questionnaire (SMEQ): There are significant differences between the OB and the BL condition in favor of the OB among the observing task, the interaction identification task, and the attention calling task ($Z = -2.941, p < 0.01$, $Z = -2.453, p < 0.05$, and $Z = -2.199, p < 0.01$, respectively). However, the OB and the BL conditions showed no differences in mental effort for the tracking task ($Z = -0.421, p = 0.674$).

Enjoyment and Level of Focus: The results showed no significant differences in both enjoyment and level of focus between the OB and the BL condition in the tracking task ($Z = -1.406, p = 0.606$ and $Z = -0.516, p = 0.614$, respectively) and the attention calling task ($Z = -2.488, p = 0.072$ and $Z = -1.715, p = 0.061$, respectively). There were significant differences in enjoyment and level of focus between the OB and the BL condition in the observing task ($Z = -2.081, p < 0.01$ and $Z = -2.825, p < 0.01$, respectively) and the interaction identification task ($Z = -1.715, p < 0.05$, and $Z = -1.992, p < 0.05$, respectively).

Social Presence and Preference: There are significant differences between the OB and the BL in favor of the OB in terms of co-presence ($Z = -4.905, p < 0.01$), attention allocation ($Z = -2.564, p < 0.01$) and perceived message understanding ($Z = -5.171, p < 0.01$). In addition, all participants chose the OB as the preferred user interface.

7 DISCUSSION

Overall, the participants rated our proposed observation system with better co-presence, attention allocation and perceived message understanding due to the fact that OB represented the students as avatars. All of the participants preferred the OB over the BL condition, as reflected by the better performances in all of the tasks. We believe that our optimization helps reduce the visual clutter in the scene so that the instructor can have a better observing experience with ObserVAR.

7.1 General feedback

Some participants made comments regarding the avatars' appearance in OB. *P3* stated that we should "make the avatars cuter", and *P12* suggested to improve the quality of the avatars as the current faceless avatars were disturbing the interaction. *P4* suggested that the avatars should be in different colors to make the task easier. *P5* suggested to change the gaze visualization to being see-through, according to *P5*, the avatars' head movements were already enough to determine the gaze location. Overall, the participants agreed that tasks can be easier if each avatar had a more distinguishable appearance.

7.2 Tracking Task

In the study, participants tracked the students accurately in both the OB and BL (90% and 83.33%). We observed that the participants had problems when the students' gaze moved to the ceiling in both conditions. When the students gazed to the ceiling, the avatars and the gaze locations could not fit within the participants' field-of-view due to the avatars were standing on the floor. Thus, the participants were required to shift their attention back and forth between the avatars and the gaze locations. Some participants adapted to this situation by focusing on the avatars and inferring the virtual objects position from the avatars' gaze. In the BL, the participants also had problems when the student's gaze moved to the ceiling. The participants sometimes did not notice when the icons were moved up to the ceiling due to their small size.

7.3 Observing Task

From the user study, we found that the OB performed better than the BL condition in all aspects including accuracy, task completion time, task difficulty, mental effort, level of enjoyment and level of focus. In the BL, the participants had trouble distinguishing the gaze of each student since the icons overlapped, even though we provided a method to highlight an icon and make the icon more distinguishable (Figure 7(b)). We also found that the participants usually missed the students' gaze in the BL, especially when the gaze was located on the ceiling or floor, since the icons are smaller compared to the avatars. In the OB, the avatars are separate from the gaze location and easily distinguishable. The participants can also easily locate the avatars because they are always standing on the floor. However, the avatars were separated from the gaze location and required more space than icons, and the participants had to move around in the physical environment to gain a better view of the avatars.

7.4 Interaction Identification Task

The quantitative results show that the participants recognized changes in the position, gesture and color of the icon. However, the participants favored the OB more than the BL because of the similar issues already indicated, i.e., the icons are hard to locate and distinguish.

7.5 Attention Calling Task

The OB performed better than the BL condition in terms of accuracy, task completion time, task difficulty and mental effort. Since there are only icons in the BL, the participants had problems identifying the students that looked at the participants because the icons were located under the participants. The participants also had difficulty locating other students' gazes in the BL, since icons are harder to spot than avatars. In the OB, the avatars helped the participants distinguish between the students that looked at the participants and the students that looked elsewhere. Nevertheless, the participants commented that being looked at by multiple avatars caused them to feel nervous.

We summarize the limitations of the proposed method based on the discussion above. First, the participants had problems when the students' were gazing at a high ceiling because the visualization cannot fit into the participants' field of view. It is suggested that

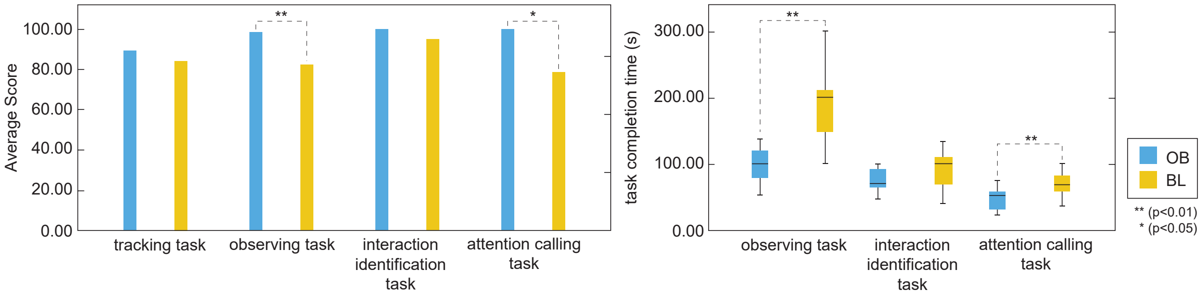

Figure 8: Accuracy and task completion time of each task.

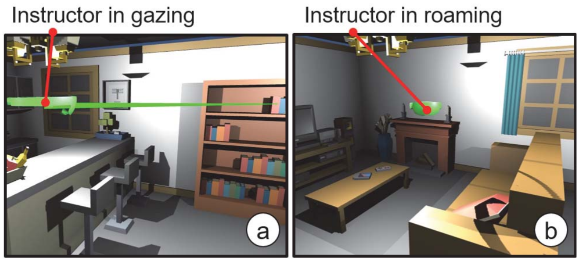

Figure 9: Two representation modes where the instructor appears in the students' VE (a) gazing mode, (b) roaming mode.

our proposed technique is not suited for VEs which feature tall objects such as skyscrapers. Second, this study only verifies the visualization setting which is suitable for observing and guiding a group of students in a static VE setting (the virtual objects are stationary). We anticipate that a dynamic VE will require additional cues such as off-screen object visualization in order to provide instructor awareness of the changing scenes. Such investigation should be explored in our future study.

8 POTENTIAL APPLICATION

We introduce a potential application based on the ObserVAR. The application allows the instructor to observe a VE classroom in multiple scenarios.

The system can potentially be used for the instructor to control the VE, choose a preferred visualization setting, and change the avatars' appearance through a control menu. As shown in the user study, the WS setting is the preferred setting to observe the entire classroom. Therefore, the WS setting was selected as a default visualization setting in the ObserVAR system. Although we use pre-recorded data in our user study, our demonstration session shows that the system allows an instructor to guide remote students in real-time. In our demonstration session, the android applications created by Unity were distributed to 5 students located in a nearby room. An instructor, in a different room, observed the avatars of the students using our system. A voice-conferencing software is used for voice communications between the instructor and the students.

The instructor can choose to observe a student in detail by focusing on that student and performing a 'select' gesture. Once selected, the FPV setting can be visualized on top of each student avatar. The VE is resized according to the instructor's hand movement where he/she can see the entire VE similar to the WIM setting. In addition, the instructor can also choose two modes to represent him/herself in the VE using either *gazing mode* or *roaming mode*. In the gazing mode, the instructor's embodiment is always shown on the students' screen to capture the attention of the students as shown in Figure 9(a). The instructor's gaze is transformed from the WS setting to match the students' VE (inverse transformation of Figure 3). In the roaming mode, since the WS setting mapped the VE to the instructor's physical environment, the instructor's physical location relative to the WS environment visualization can be transformed to position the instructor inside the students' VE. Thus, the instructor can show his/her physical movement inside the students' VE. Figure 9(b) shows an instructor who is standing between the virtual sofa and virtual table in the student's VE.

9 LIMITATION AND FUTURE WORK

We believe that ObserVAR can improve the observation performance by using different avatar appearances for each student. However, to clarify all interaction aspects between an instructor and the students, a formal user study in complex classroom situations will be conducted to further improve guiding tasks in VR classrooms. The current system assumes that the students are in remote locations. If they are co-located with the instructor, we expect that the visualizations and visual clutter issues for the co-located students would require further study. In addition, the participants in our experiments were university students. Students from other levels of education might behave differently in the VR classroom.

10 CONCLUSION

In this paper, we have presented ObserVAR, an augmented reality approach, as an observation tool that allows instructors to observe their students in VR classrooms. First, we investigated the visualization settings, which are suitable for instructors to observe and instruct students in a VE. We designed and determined three visualization settings for observing multiple users, namely, first person view, world in miniature, and world scale. Our preliminary study showed that the world scale visualization setting can improve the awareness of the instructor. We then further improved the visualization setting by reducing the visual clutter on the screen by optimizing the students' avatar positions using a force-directed graph drawing algorithm. The proposed system was compared with existing user interfaces in our formal user study to evaluate its advantages and limitations. ObserVAR was found to improve an instructor's observation experience, awareness, and social-presence under different scenarios compared to existing VE classroom technologies.

ACKNOWLEDGMENTS

This project was supported in part by Hong Kong RGC's Theme-based Research Scheme (grant no. T44-707/16-N), the UGC Teaching and Learning Funding Scheme (grant number F0667), the Innovation and Technology Fund (grant no. ITS/227/17FP and ITS/388/17FP), and the Osaka University Researcher Development Fund (grant no. J18110201).

REFERENCES

[1] B. Avery, C. Sandor, and B. H. Thomas. Improving spatial perception for augmented reality x-ray vision. In *Proceedings of the 2009 IEEE Virtual Reality Conference*, pages 79–82, March 2009.

[2] R. Bane and T. Hollerer. Interactive tools for virtual x-ray vision in mobile augmented reality. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '04, pages 231–239, Washington, DC, USA, 2004. IEEE Computer Society.

[3] G. D. Battista, P. Eades, R. Tamassia, and I. G. Tollis. *Graph Drawing: Algorithms for the Visualization of Graphs*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1st edition, 1998.

[4] E. Bozgeyikli, A. Raij, S. Katkoori, and R. Dubey. Point & teleport locomotion technique for virtual reality. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '16, pages 205–216, New York, NY, USA, 2016. ACM.

[5] B. Brown, I. MacColl, M. Chalmers, A. Galani, C. Randell, and A. Steed. Lessons from the lighthouse: Collaboration in a shared mixed reality system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 577–584, New York, NY, USA, 2003. ACM.

[6] S. Charara. Over two million school kids have gone on Google Expeditions. `https://www.wareable.com/vr/expeditions-two-million-cardboard-6665`, 2017. [Online; accessed 12-June-2019].

[7] K. R. Dillman, T. T. H. Mok, A. Tang, L. Oehlberg, and A. Mitchell. A visual interaction cue framework from video game environments for augmented reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 140:1–140:12, New York, NY, USA, 2018. ACM.

[8] P. Eades and H. M. Lin. *Navigating Clustered Graphs Using Force-Directed Methods*, pages 191–215. 2004.

[9] O. Erat, W. A. Isop, D. Kalkofen, and D. Schmalstieg. Drone-augmented human vision: Exocentric control for drones exploring hidden areas. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1437–1446, April 2018.

[10] T. M. J. Fruchterman and E. M. Reingold. Graph Drawing by Force-directed Placement. *Software-Practice and Expererience*, 21(November 1991):1129–1164, 1991.

[11] Google. Google expedition. `https://edu.google.com/products/vr-ar/expeditions/?modal_active=none`. Accessed: 2018-11-27.

[12] R. Grasset, P. Lamb, and M. Billinghurst. Evaluation of mixed-space collaboration. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '05, pages 90–99, Washington, DC, USA, 2005. IEEE Computer Society.

[13] J. Grubert, T. Langlotz, S. Zollmann, and H. Regenbrecht. Towards pervasive augmented reality: Context-awareness in augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 23(6):1706–1724, June 2017.

[14] A. Grundhöfer and D. Iwai. Recent advances in projection mapping algorithms, hardware and applications. *Computer Graphics Forum*, 37(2):653–675, 2018.

[15] J. Gugenheimer, E. Stemasov, J. Frommel, and E. Rukzio. Sharevr: Enabling co-located experiences for virtual reality between hmd and non-hmd users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 4021–4033, New York, NY, USA, 2017. ACM.

[16] P. C. Harms and P. F. Biocca. Internal consistency and reliability of the networked minds measure of social presence, 2004.

[17] H. Ibayashi, Y. Sugiura, D. Sakamoto, N. Miyata, M. Tada, T. Okuma, T. Kurata, M. Mochimaru, and T. Igarashi. Dollhouse vr: A multi-view, multi-user collaborative design workspace with vr technology. In *SIGGRAPH Asia 2015 Emerging Technologies*, SA '15, pages 8:1–8:2, New York, NY, USA, 2015. ACM.

[18] S. Julier, Y. Baillot, D. Brown, and M. Lanzagorta. Information filtering for mobile augmented reality. *IEEE Computer Graphics and Applications*, 22(5):12–15, Sept 2002.

[19] S. Kasahara and J. Rekimoto. Jackin: Integrating first-person view with out-of-body vision generation for human-human augmentation. In *Proceedings of the 5th Augmented Human International Conference*, AH '14, pages 46:1–46:8, New York, NY, USA, 2014. ACM.

[20] K. Kim, M. Billinghurst, G. Bruder, H. B. Duh, and G. F. Welch. Revisiting trends in augmented reality research: A review of the 2nd decade of ismar (2008–2017). *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2947–2962, Nov 2018.

[21] S. G. Kobourov. Force-directed drawing algorithms. 2004.

[22] J. Lehikoinen and R. Suomela. Accessing context in wearable computers. *Personal Ubiquitous Comput.*, 6(1):64–74, Jan. 2002.

[23] Microsoft. Microsoft HoloLens spatial mapping. `https://docs.microsoft.com/en-us/windows/mixed-reality/spatial-mapping`. Accessed: 2018-11-22.

[24] Z. Noh, M. S. Sunar, and Z. Pan. A review on augmented reality for virtual heritage system. In M. Chang, R. Kuo, Kinshuk, G.-D. Chen, and M. Hirose, editors, *Learning by Playing. Game-based Education System Design and Development*, pages 50–61, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[25] Oculus. Rift core 2.0. `https://www.oculus.com/blog/welcome-to-rift-core-beta-now-available/`. Accessed: 2019-02-27.

[26] J. S. Pierce, B. C. Stearns, and R. Pausch. Voodoo dolls: Seamless interaction at multiple scales in virtual environments. In *Proceedings of the 1999 Symposium on Interactive 3D Graphics*, I3D '99, pages 141–145, New York, NY, USA, 1999. ACM.

[27] T. Piumsomboon, A. Dey, B. Ens, G. Lee, and M. Billinghurst. The effects of sharing awareness cues in collaborative mixed reality. *Frontiers in Robotics and AI*, 6:5, 2019.

[28] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas, and M. Billinghurst. Mini-me: An adaptive avatar for mixed reality remote collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 46:1–46:13, New York, NY, USA, 2018. ACM.

[29] H. C. Purchase, R. F. Cohen, and M. James. Validating graph drawing aesthetics. In F. J. Brandenburg, editor, *Graph Drawing*, pages 435–446, Berlin, Heidelberg, 1996. Springer Berlin Heidelberg.

[30] G. Reitmayr and D. Schmalstieg. Scalable techniques for collaborative outdoor augmented reality. In *Proceedings of the 3rd IEEE and ACM international symposium on mixed and augmented reality*, ISMAR'04, pages 1–10, 2004.

[31] J. Sauro and J. S. Dumas. Comparison of three one-question, post-task usability questionnaires. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1599–1608. ACM, 2009.

[32] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *The Craft of Information Visualization*, pages 364–371. Elsevier, 2003.

[33] R. Stoakley, M. J. Conway, and R. Pausch. Virtual reality on a wim: Interactive worlds in miniature. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '95, pages 265–272, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co.

[34] S. Thanyadit, P. Punpongsanon, and T.-C. Pong. Efficient information sharing techniques between workers of heterogeneous tasks in 3d cve. *Proc. ACM Hum.-Comput. Interact.*, 2(CSCW):172:1–172:19, Nov. 2018.

[35] K.-T. Yang, C.-H. Wang, and L. Chan. Sharespace: Facilitating shared use of the physical space by both vr head-mounted display and external users. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, UIST '18, pages 499–509, New York, NY, USA, 2018. ACM.

[36] S. Yuki, H. Keita, Y. Ryo, K. Rie, and S. Yoichi. Browsing group first-person videos with 3d visualization. In *Proceedings of the ACM International Conference on Interactive Surfaces and Spaces*, ISS '18, pages 55–60, New York, NY, USA, 2018. ACM.

[37] F. R. H. Zijlstra and L. Van Doorn. The construction of a scale to measure subjective effort. *Delft, Netherlands*, 43, 1985.