ARGarden: 3D outdoor landscape design using handheld augmented reality with multi-user interaction

To cite this article: F E Fadzli *et al* 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* **979** 012001

IOP Conf. Series: Materials Science and Engineering **979** (2020) 012001 doi:10.1088/1757-899X/979/1/012001

ARGarden: 3D outdoor landscape design using handheld augmented reality with multi-user interaction

F E Fadzli [1,2] *, M A Mohd Yusof [1,2], A W Ismail [1,2], M S Hj Salam [1,2] and N A Ismail [1,2]

[1] Mixed and Virtual Environment Research Lab (mivielab), ViCubeLab, Universiti Teknologi Malaysia, 81310 Johore, Malaysia
[2] School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 Johor, Malaysia

*fazliaty.edora@gmail.com

Abstract. Augmented Reality (AR) is a technology to display the virtual object by overlaying on the physical marker in the real environment. Nowadays, the handheld device becomes a trend in AR since it has high processing power and compatible with the AR system. However, most of the AR applications did not fully support the interaction from multiple users for the collaborative AR interface. Therefore, this research aim is to develop multi-user interaction for collaborative handheld AR. There are a few phases involved to design and develop the collaborative AR application for the handheld device called ARGarden. All the interaction that has been done for the virtual object transformation such as translation, rotation and scaling can be seen by the other users, and this data is synchronized by using a network protocol. Then, a queue system is implemented for the multi-user interaction to allow the virtual object can only be selected by one user at one time, and the other users must wait for their turn. The significance of this research is all the method that has been implemented for the ARGarden can be applied in different fields such as simulation, modelling, medical and education.

1. Introduction

Virtual reality (VR) or virtual environment is an intuitive way to enable the participant to accomplish the task in the virtual world which impossible in the real world by creating the synthesis of reality [1]. This technology makes the user feel they self-inside the VR, and this is called immersive. Immersive happen when there is realism, low latency which is, the time taken for refresh is around 17 milliseconds and smooth, which the refresh rate is about 80 to 120 Hz and framerate per second is 60 fps for GPU and 30 fps for CPU. However, the purpose of augmented reality (AR) is the opposite of VR. VR is used to create the virtual world and replace reality, but AR is used to enhance reality. In order to create mixed reality (MR) in real-time, AR technology refers to placing the virtual elements in the actual real environments view to improve and enhance the senses of human perceptions in the real world [2]. Therefore, AR can improve user interaction with the real world by displaying important information that cannot be recognized or identified directly by using their senses. This important information can be in virtual, for example, mobile augmented reality shopping application for Dulux Paint Expert which enables the user to see the realistic Dulux Trade paint colours appear on the wall indirectly give them the paint information [3].

Nowadays, many prefer to use AR because it helps them to understand some information much better rather than look at the direct view of the information. For example, the usage of VR and AR in dentistry not only uses the system for learning and teaching from their perspective but also training their skill and improve the hand-eye coordinate indirectly help the user to correct the posture and skill [4]. There is also an AR application for modelling 3D object in real environment [5].

In AR, there are a few problems that are still unsolved until today. Most of the problem usually related to the fundamental of AR, which consists of four categories, display technique, tracking technique, interface, and interaction. To enable the user to seamlessly interact with the virtual object by implementing the most suitable interaction technique for AR application is one of the essential research in AR [6]. AR interface often only for the view, and the users cannot manipulate. Then, the interaction comes. Instead of only can browse the AR by the user, with the interaction the user can manipulate the virtual element in the AR. The interaction must be intuitive and be natural as possible to enhance the natural user interface (NUI). Besides that, there are four techniques to display AR, such as a projector, depth-HD camera, webcam, and mobile. A mobile device is also known as a handheld device. Therefore, each of the display techniques must have a different type of suitable approach for the interaction. In addition, there are four types of interface in AR, tangible interface (TUI), Transitional Interface, collaborative AR interface, and hybrid. By using a collaborative AR interface, the AR view becomes collaborative by connecting between two or more users through the internet connection [7].

Mobile AR has ubiquitous availability, strong computing power and by relating physical-world objects and virtual information makes the mobile AR can support the situated or immersive perception of a complex concept [9]. Besides that, interaction is one of the critical issues in AR in terms of accuracy and robustness. It is because implementing an appropriate interaction technique can affect the user experience in AR. The 3D object manipulation has been a challenge for the interface designer since it involves the control of six degrees of freedom (6DOF) [10]. Furthermore, to improve the communication between the user in AR is by making the AR become collaborative. Collaboration in AR can improve the users' awareness and give attention to their surroundings with the local user [8].

2. Handheld AR Application

The handheld device is used as a display technique. Before it can be used, a few configurations need to be made. The right API is required to run the application on the handheld device without any problem or error. Besides that, the handheld device also comes with a built-in camera. The camera is used to video stream the target marker. If the target marker is recognized as a registered marker, then the AR 3D Model appeared on the device display screen at the center of the marker. Furthermore, the handheld device also comes with a touchscreen screen.

2.1. Methodology

The first phase describes the user interaction within the collaborative AR interface environment. It is crucial to identify the remaining issues in handheld AR and the fundamental of AR. It also includes the touchscreen interaction for the handheld device. The findings of this phase have been further discussed in the literature review.

The second phase is to develop an AR application in a handheld device and to design AR collaborative interface. At this phase, the logic is being coded, and the AR view is enabled. After that, user interaction with more than one user, a multi-user interaction has been created to actualize collaborative AR interface Multi-user interaction applied touch gesture for interaction and to allow collaborative works, an internet protocol is determined. In order to enable AR technology in a handheld device, a natural feature tracking has been executed in this phase. ARGarden application has been developed with the capability to add more than one user to interact with the 3D contents.

The third phase is to integrate ARGarden application with multi-user interaction for collaborative AR. At this phase, photon service is used to enable two or more users to connect with each other during collaborative AR. Figure 1 shows the methodology of this research.

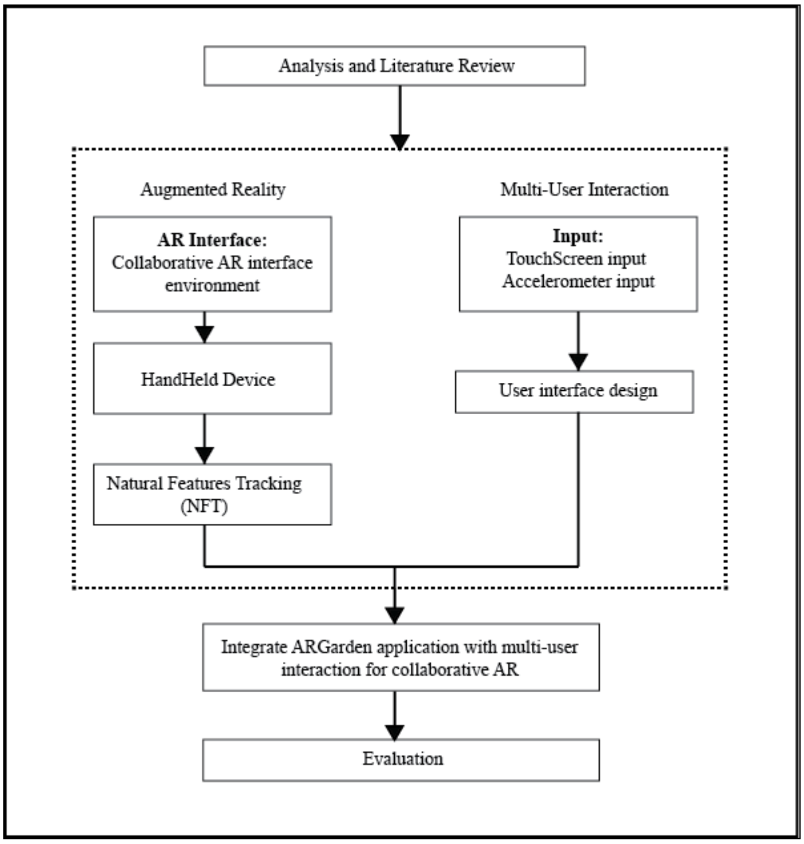


Figure 1. Research Methodology.

2.2. AR Handheld Interface

The features-based is a tracking technique that detects the feature on a real image by using a camera. Figure 2 shows our RGB marker has been converted into features. We used a single marker, coloured image in RGB as our image target and uploaded it into the database in the Vuforia SDK. This features-based tracking system will convert our image target from RGB to grayscale version to identify the features that can be used for recognition and tracking.

Figure 2 (a) explains the process of converting an RGB image into features. The marker tracking process has been used in this research; a printed image with RGB colour is tracked by a camera, then the image is converted into grayscale as in figure 2 (b), and with the default threshold values, the image has produced the features, as in figure 2 (c). The features are generated and stored in a dataset. The marker with the unique pattern is recognized by their features; once the camera captures the image, the dataset is retrieved. The features are used to register an object. Once the camera detects the image with the assigned object, the object is overlaid on the top of the image.

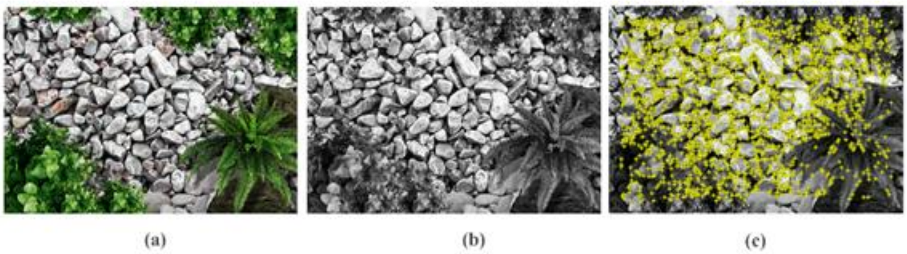


Figure 2. Natural Features Tracking (NFT) for AR Tracking.

IOP Conf. Series: Materials Science and Engineering **979** (2020) 012001 doi:10.1088/1757-899X/979/1/012001

Figure 3 presents the AR interface architecture for this research. The virtual element contains the data of the 3D object. This data is used to spawn the object in the AR scene in format. fbx. This 3D data has the 3D geometry and the texture of the object. Besides that, the tracking technologies used feature tracking, which is feature-based tracking. Besides that, the interaction used is multi-user interaction for a handheld device. The interaction that is involved in 3D object manipulation is positioning and orientation. The result from this interaction was calculated and displayed through the handheld display.

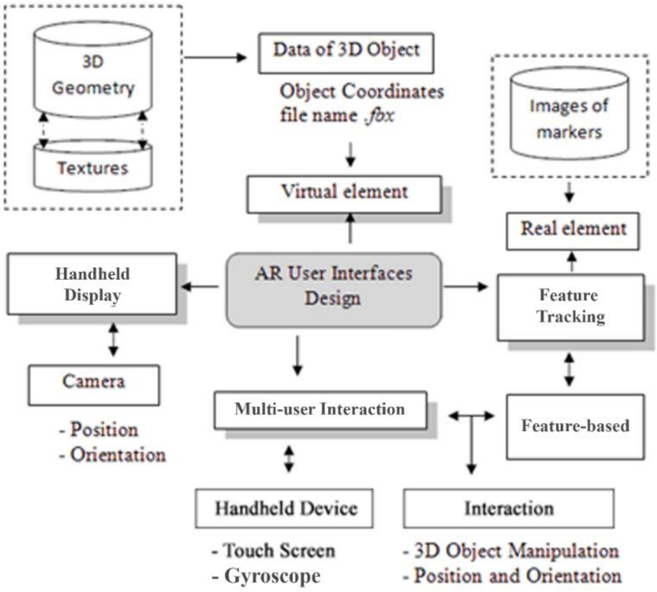


Figure 3. AR User Interface Design with Handheld.

2.3. Touch-Screen and Gyroscope Input

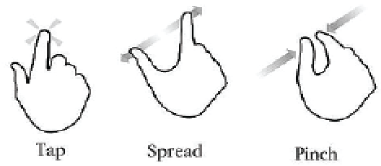
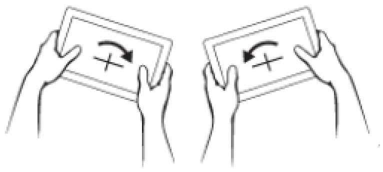
The input used in this research is touchscreen input and gyroscope input, as shown in Fig. 3. Before the device tilting input can be received, the object needs to be selected first by using touchscreen input. If the touch input receives, then the raycast is generated, and it checked if the raycast is touching any object. When the object is touch, it gives feedback to the user by highlighting the object and creating a line between the user device and the selected object. Then, any translation of the object can be done by swiping the touch input on the screen. Besides that, the accelerometer is a technology that is used to measure tilt angle, rotation, and it is applied in changing the device screen orientation vertically or horizontally. For this research, the accelerometer is used to manipulate the object rotation by tilting the device. After any translation has been done, the viewpoint is calculated before the result display on the screen.

Table 1 shows the touchscreen gesture that has been used to get input from the user. (a) Tap gestures are used to instantiate the object. Tap is a gesture by using one index finger to touch on the screen and the gesture to scale up and scale down the object by using (b) spread and (c) pinch. Spread is a gesture by spreading the thumb finger and index finger apart while touching on the screen index while the pinch is a gesture by pinching the thumb finger and index finger together while touching on the screen.

When the object is touch, selected or deselected, the touched object is sent through the network, and it tells all users the state of the object. If the user is the owner of the object, then the object can be manipulated by the user to do the transformation. Table 1 shows the tilting gesture when doing the rotation transformation. When the user tilts the device to the right (a), the object is rotated to the right, and when the user rotated to the left, as shown in (b), the object is turned to the left. The tilt input is received from gyroscope input.

IOP Conf. Series: Materials Science and Engineering **979** (2020) 012001 doi:10.1088/1757-899X/979/1/012001

Table 1. Interaction using touchscreen and gyroscope.

Touch-screen gesture	Tilting gesture
 (a) (b) (c)	 (a) (b)

3. Multi-user Interaction

The handheld device also comes with a built-in camera. This camera is used to video stream the target marker. If the target marker is recognized as a registered marker, then the AR 3D Model appeared on the device display screen at the centre of the marker. Furthermore, the handheld device also comes with a touchscreen screen. Input that is used in this research is touchscreen input and gyroscope input. Before the gyroscope input can be received, the object needs to be selected first by using touchscreen input. If the touch input receives, then the raycast is generated, and it checked if the raycast is touching any object. When the object is touch, it gives feedback to the user by highlighting the object and creating a line between the user device and the selected object.

Before this stage, the application was working only for a single user. Since the AR need to be collaborative, the user needs to be connected with each other through the network connection. In order to enable this environment, Photon Unity Networking (PUN) services were used, and figure 4 shows the flow process for multi-user.

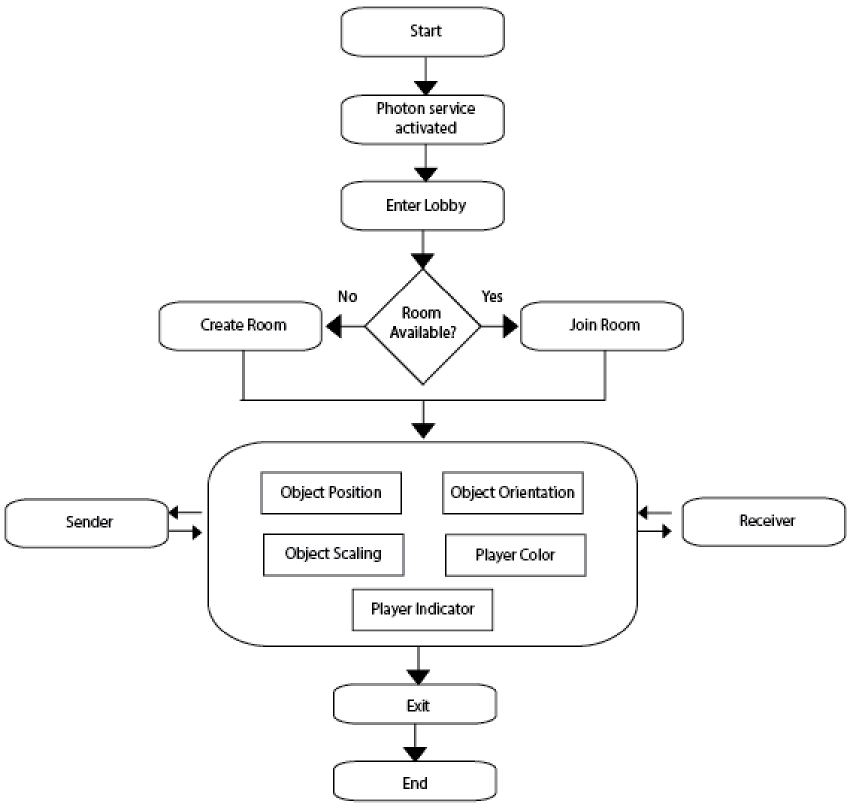
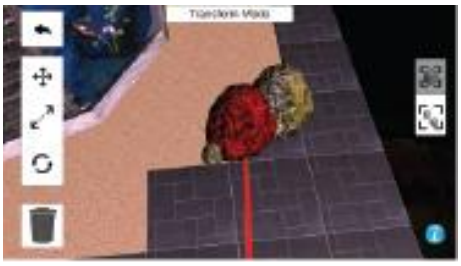


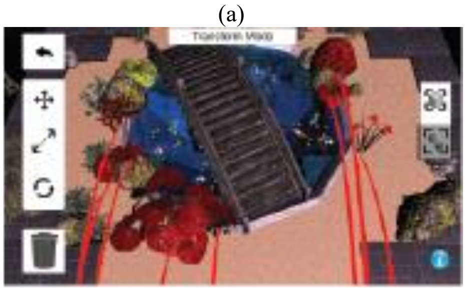
Figure 4. Multi-user Interaction in AR.

First, to activate and use the photon services, the photon app id needs to be created on its website, and this id is used for the photon configuration. After the service is activated, the user can enter the lobby. Then the user can create a room for the other user to join it. The user is responsible for creating the room is assigned as the master client. After the users connected in a room, process sender and receiver happened through the transmission control protocol (TCP). During this process, the data that is sent and received by the user is the object position, object orientation, object scaling, player colour, and player indicator. The object position, orientation and scaling were manipulated by using touchscreen input and gyroscope input.

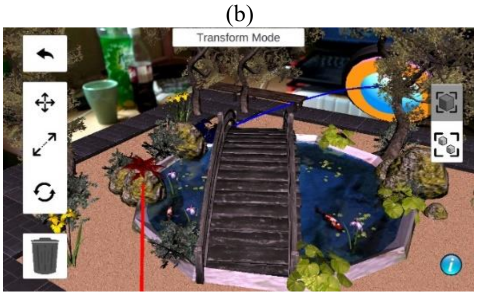
There are two types of selection which are single selection and multiple selections. The implementation of a single selection is shown in figure 5 (a), and multiple selections are shown in figure 5 (b). Touch-screen and gyroscope input were used to enable the user to interact with the object in the AR environment. The user can move the object by dragging their hand on the screen. The gyroscope is used for the user to rotate the object while touching the screen. The indicator is spawned when the object was selected as in figure 5 (c). This indicator also can be seen by the other user in blue colour.



(a)



(b)



(c)

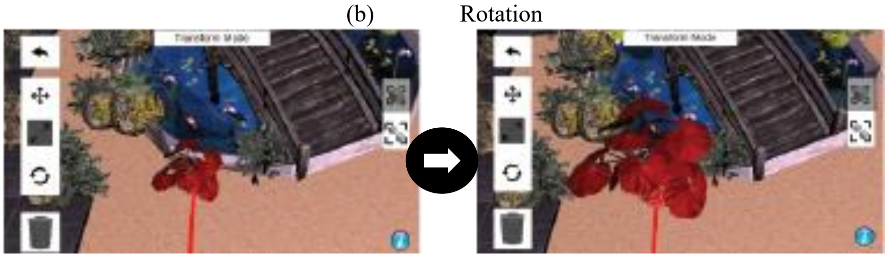
Figure 5. (a) single selection, (b) multiple selection, (c) multi-user selection.

Next, a few transformations have been applied, which are translation, rotation, and scaling. Figure 6 (a) shows the translation and figure 6 (b) shows the rotation. Figure 6 (c) shows the scaling and figure 6 (d) shows the uniform scaling, to perform scaling for the multiple object selection.

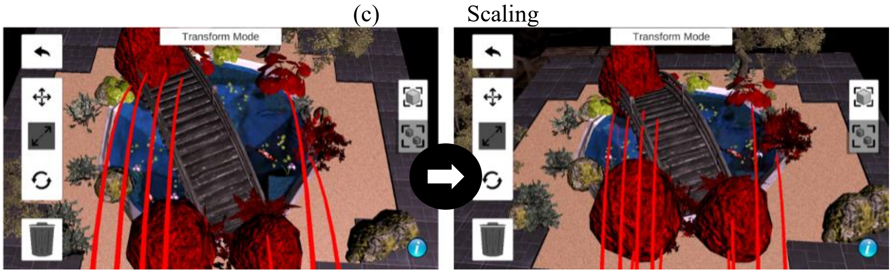
IOP Conf. Series: Materials Science and Engineering **979** (2020) 012001 doi:10.1088/1757-899X/979/1/012001

(a) Translation

(b) Rotation



(c) Scaling



(d) Uniform scaling

Figure 6. Manipulation in ARGarden application.

4. 3D Outdoor Landscape Design

ARGarden application is a prototype to implement method multi-user interaction for collaborative handheld AR. In this application, there are a few features and options that have been implemented. First are the features-based tracking technique. Even a part of the marker has been occluded, and the AR still can be displayed because another part of the marker can still be recognized by using feature-based tracking. In addition, in Vuforia, they provided a feature called extended tracking. This feature has been implemented in this application. Extended tracking is a concept that enables the target pose information to be still available even the target is no longer in the camera field of view. Therefore, by using this feature, it can improve the tracking performance and continue tracking even the target no longer in view.

Next, the network service is used in this application to create AR collaborative environment. This environment every object in the scene to be synchronized by the server. When it is involved multiple users, a queuing system has been implemented. Therefore, by using the queuing system, the user cannot touch or select any object that is currently selected by the other user. The user was also given a notification to tell them that they are already connected to the lobby. Once the users enter the application, the first screen has appeared as in figure 7 (a). The user was also provided with a tutorial when the first time they were using the application. The transform mode interface has appeared when `EditWorld` button is clicked. `Camera` button is used to take the screen-shot of the application, and UI is disabled during the screen-shot is taken. The screen-shot picture is saved in format .png. `ItemList` button allows the user to choose any type of object that they want to instantiate. `Information` button is used to display the tutorial as a guide for the user. Figure 7 (b) shows three users log-in into ARGarden.

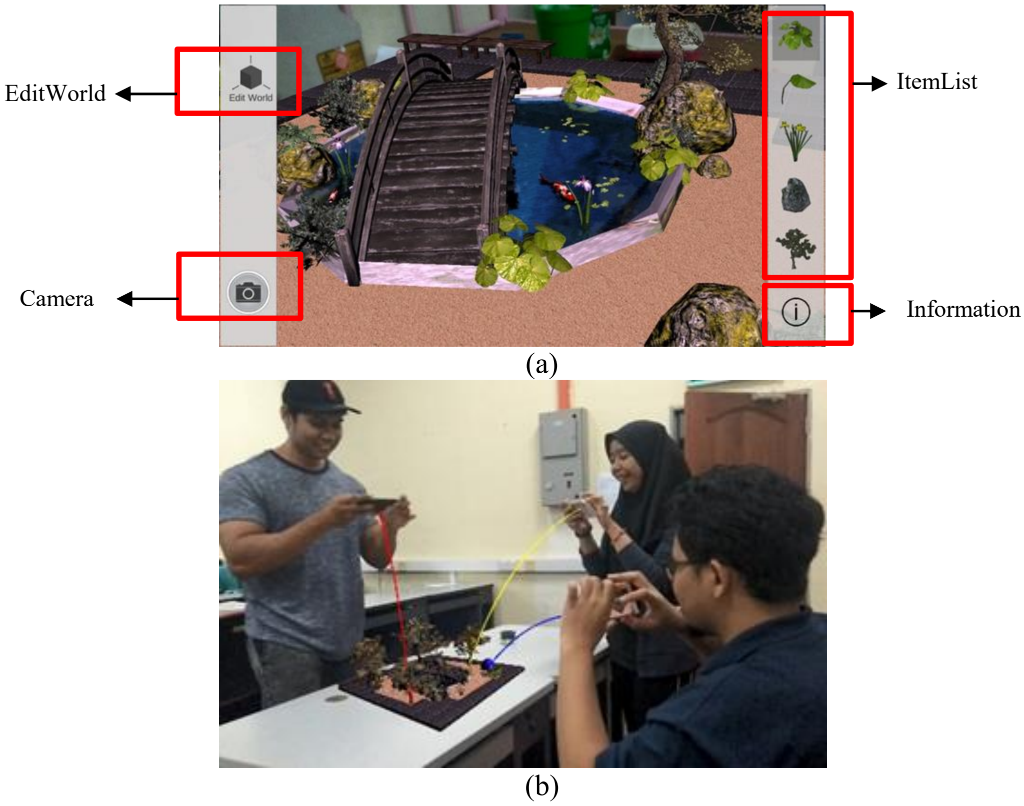


Figure 7. Multi-user Interaction in ARGarden Handheld AR.

Once `EditWorld` is selected, the user is able to modify the 3D outdoor landscape. The `Back` button is used to allow the user to back to the main menu. The `Deselect` button also deselects all currently selected objects when transform mode changes to the main menu. The `Translate` button allows the user to translate the selected object and the `Scale` button allows the user to scale up or scale down the selected object. The `Rotate` button allows the user to rotate the selected object, and the Select button allows the user to select only one object at one time. Lastly, `Multi` button allows the user to select more than one object at one time, and the `Delete` button allows the user to delete the selected object. All available buttons have been assigned as in figure 8.

IOP Conf. Series: Materials Science and Engineering **979** (2020) 012001 doi:10.1088/1757-899X/979/1/012001

Figure 8. Transform Mode in `EditWorld` option.

5. Evaluation

In the experiment setup, the physical printed A3 marker was placed on the table where the users can move around the table to scan the marker. Two users with the handheld device are needed to create a collaborative multi-user interaction environment. Each of the handheld device used has a built-in camera that is used to track the marker, and the devices have a feature to receive the touchscreen input for the user to interact with the virtual object. Then, all the handheld device used also has an internet connection that is used to connect the users by using the network.

The participant age is between 22 and 23 years old, with a total of 10 students. They were given an instruction manual on how to use the prototype before working on their own. Then they were asked to fill the information form that asking them if they have any experience with AR and in using AR that is more than one user. Then, they were given time at least 4 min to interact with the 3D virtual object in the collaborative AR environment. Then the questionnaire and respondents' comments were used to collect the data. The pre-questionnaire asked the respondent about their experience in using AR and their experience in using AR with multiple users. Post-questionnaire asked the respondent about the feedback upon using this application.

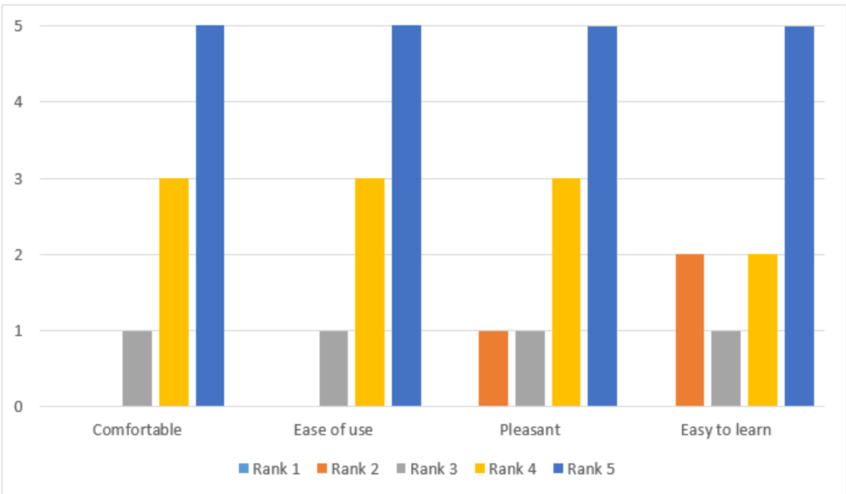


Figure 9. User experience with ARGarden.

The evaluation was guided referred to ARPad [14]. Then the questionnaire and respondents' comments were used to collect the data. The rank provides 5 levels, 1 indicates the lowest and 5 indicates the highest. The evaluation of ARGarden is to measure user experience in terms of comfortable, ease of use, pleasant and easy to learn. Figure 9 shows in comfortable user to the ARGarden application, 6 respondents have assigned Rank 5 and 3 respondents assigned moderate with Rank 3. In terms of ease of use, the results show the same to the comfortable. The users feel pleasant with the application, and however, in terms of easy to learn, the users feel joining a room to invite multi-user is hard when the task interrupts with the internet connection. A stable network is required to work with the ARGarden application. User feedback is positive on the colour indicator as visual feedback to differentiate the users in ARGarden. For future work, there will be more evaluation needed to measure the application on how touchscreen and gyroscope are able to improve the user experience when using the application.

6. Conclusion

This paper has discussed in detail 3D outdoor landscape design using handheld AR with multi-user interaction implemented in ARGarden application. A few limitations, however, were identified in this research which one of it is the internet connection must be fully functional for the collaborative AR part to work. Without an internet connection, the users cannot connect with each other and the AR view cannot synchronize. Besides that, the object also cannot update its position correctly if the internet connection was unstable or slow.

As for future work, there are a few recommendations that can be applied to improve the user experience when using the application. First, the marker-based tracking can be changed into marker-less tracking. Therefore, the user can use the application anywhere without needing a physical printed marker. They can just scan a ground plane and set the area they want to display the AR. Besides, for the interaction, it was recommended to try to use the real hand gesture to enable the user to interact with the virtual object just like the real world.

Acknowledgements

We would like to express our appreciation to Mixed and Virtual Reality Laboratory (mivielab), ViCubeLab at Universiti Teknologi Malaysia (UTM) for the facilities and technical supports.

References

[1] L V Zhihan, L Xiaoming, and L Wenbin 2017 Virtual reality geographical interactive scene semantics research for immersive geography learning *Neurocomputing* 254 71-78
[2] J M Mota, I Ruiz-Rube, J M Dodero, and I Arnedillo-Sánchez 2018 Augmented reality mobile app development for all *Computers & Electrical Engineering* 65 250-260.
[3] S G Dacko 2017 Enabling smart retail settings via mobile augmented reality shopping apps *Technological Forecasting and Social Change* 124 243-256.
[4] T K Huang, C H Yang, Y H Hsieh, J C Wang, and C Hung 2018 Augmented reality (AR) and virtual reality (VR) applied in dentistry *The Kaohsiung journal of medical sciences* 34(4) pp 243-248.
[5] F E Fadzli, and A W Ismail 2019 VoxAR: 3D Modelling Editor Using Real Hands Gesture for Augmented Reality *In 2019 IEEE 7th Conf. on Systems, Process and Control (ICSPC)* pp 242-247
[6] A W Ismail, M Billinghurst, and M S Sunar 2015 Vision-Based Technique and Issues for Multimodal Interaction in Augmented Reality. *In Proc. of the 8th International Symp. on Visual Information Communication and Interaction* pp 75-82
[7] F Zhou, H B L Duh, and M Billinghurst 2008 Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. *In Proc. of the 7th IEEE/ACM Int. Symp. on Mixed and Augmented Reality* pp 193-202
[8] D Datcu, S G Lukosch, and H K Lukosch 2014 A collaborative game to study the perception of presence during virtual co-location *In Proc. of the companion publication of the 17th ACM conf. on Computer supported cooperative work & social computing* 5-8
[9] F Ke, and Y C Hsu 2015 Mobile augmented-reality artifact creation as a component of mobile computer-supported collaborative learning *The Internet and Higher Education* 26 33-41.

IOP Conf. Series: Materials Science and Engineering **979** (2020) 012001 doi:10.1088/1757-899X/979/1/012001

[10] A W Ismail. *User Interaction Technique With 3D Object Manipulation in Augmented Reality Environment* 2011 (Doctoral dissertation, Universiti Teknologi Malaysia)
[11] A W Ismail, Noh Z Ismail, M S Sunar, and M Najib 2008 Augmented Reality Theory and Applications. *Advances in Computer Graphics and Virtual Environment* 88-105.
[12] P Milgram and H Colquhoun 1999 A taxonomy of real and virtual world display integration. Mixed reality: Merging real and virtual world (vo1, pp 1-26.)
[13] A W Ismail, and M S Sunar .2009 Collaborative augmented reality: multi-user interaction in urban simulation *In Int. Visual Informatics Conf.* pp 382-391
[14] D Mogilev, K Kiyokawa, M Billinghurst and J Pair 2002 Ar pad: An interface for face-to-face ar collaboration *In CHI'02 extended abstracts on Human factors in computing systems* 654-655