Target selection method on the occluded and distant object in handheld augmented reality

Ajune Wanis Ismail[1], Nur Ameerah Abdul Halim[1], Rohaya Talib[2], Ahmad Johari Sihes[2]
[1]Mixed and Virtual Environment Research Lab, Faculty of Computing, Universiti Teknologi Malaysia, Johor, Malaysia
[2]Faculty of Social Sciences and Humanities, Universiti Teknologi Malaysia, Johor, Malaysia

Article Info

Article history:

Received Jul 5, 2022
Revised Oct 20, 2022
Accepted Nov 1, 2022

Keywords:

Augmented reality
Distant object
Hand gesture
Occluded object
Raycasting
Target selection

ABSTRACT

Existing interaction techniques within handheld augmented reality (AR) have frequently used touchscreen input (pure two-dimensional (2D) pointing and clicking) from the handheld device's display for target selection on the virtual object. However, performing accurate target selection on a distant target object becomes challenging as the target object will appear smaller when the distance increases. Aside from that, the difficulty increases in performing target selection when another virtual object obscures the distant virtual object. Therefore, this study aims to present a target selection method to perform the target selection. We enable the raycasting technique with real hand gesture for the target selection method on the occluded and distant object in handheld AR. The leap motion device is mounted at the back of the handheld device to track the real hand gesture. The markerless tracking technology of simultaneous localization and mapping (SLAM) is implemented to enable the AR environment. Based on the results, the aim of this study was achieved.

Corresponding Author:

Ajune Wanis Ismail
Mixed and Virtual Environment Research Lab, Faculty of Computing, Universiti Teknologi Malaysia
81310 Skudai, Johor
Email: ajune@utm.my

1. INTRODUCTION

Augmented reality (AR) technology refers to inserting the virtual components into the actual real-world surroundings views which improve and enhance the senses of human perceptions in the real world [1]. Although at the early stage, AR was mainly developed for head-mounted displays (HMD) and desktop computers, with the advancement of the current technologies, handheld devices such as tablets and smartphones are preferred for interaction in AR due to their availability [2]-[4]. Despite marker-based tracking being utilized in many AR applications, recent research shows that markerless tracking techniques such as simultaneous localization and mapping (SLAM) are being opted instead for handheld AR. It enables the user to track the pose of the device, without requiring the user to set up fiducial markers or scan the environment beforehand [5].

In regard to that, one of the most essential aspects of AR applications is the interaction among the user and the virtual content within the space, as stated by [6]. Target selection permits the user to get or specify a target item in order to accomplish impending interactions on or with it [7] in which manipulation tasks are often followed by (and reliant on) selection tasks. As a consequence, poorly crafted selection procedures often impair overall user efficiency. Therefore, the target selection method is our main focus. This study addresses the issue of target selection, on the occluded and distant object in a handheld AR environment. This issue is important and often occurs during the target object's interaction. Despite the issues commonly discussed in the virtual reality (VR) environment, very less attention is taken to the issues to be discussed in AR environment.

Despite that, as AR is able to obey the VR principle, the target selection method proposed for VR environment is also applicable to be utilized in the AR environment, due to the nature of the target selection. Target selection on the occluded and distant object is often highlighted in previous research as an ongoing issue [8]-[11] when interacting with the object. In the case of selection on the occluded object (cluttered scene) either fully or partially occluded, the traditional methods may begin to lose their function in terms of precision and speed [12]. Meanwhile, the occluded target object being in distance from the user increases the difficulty for the target selection task to complete [9], [13], [14].

Despite these issues, the problem becomes worse as the conventional way of interaction for handheld devices of the touch input is often limited by the device's physical boundary and usability suffers as the on-screen content becomes occluded by the finger during interaction [15]. Interaction using a paddle-based, controller, or joystick are among the other available method for interaction [16], [17]. However, the user usually needs time to adapt to the devices in order to perform a more accurate selection. As stated by Ismail and Sunar [18], there is a bottleneck that happens, which depends on user engagement due to the interaction's artificial nature. In the hopes of removing this bottleneck, people began to study human forms such as gaze, speech, and gesture recognition to develop ways of communicating with machines naturally and intuitively [19]. Meanwhile, states that by using gesture, the interaction of natural, efficient, and intuitive activities may be achieved [2]. Therefore, we further discuss the approaches of previous researchers to attend to the highlighted issues above in the next section.

2. RELATED WORKS

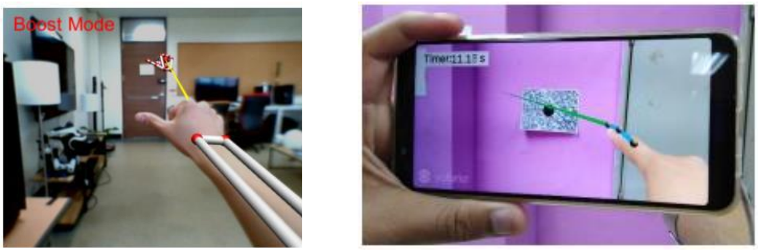
AR is being explored more deeply, enabling many upgraded and new ways of selection to be explored and implemented by researchers to provide more precise and engaging user interaction. Table 1 shows the comparison of recent approaches taken by researchers in addressing the issue of selection on the target object. As stated by Poupyrev *et al*. [20], there are two categories of selection techniques which are exocentric and egocentric. In the exocentric approach, the interaction is done by the user outside of the virtual environment. Exocentric approach can be further categorized into two categories which are world-in-miniature (WIM) and automatic scaling. Meanwhile, the egocentric approach is when the user operates directly from the virtual environment's interior as if it were a part of it. However, since this approach is often employed for the precise manipulation of the object, it is less suitable to implement for large-scale manipulation tasks. Virtual hand and virtual pointer are two types of egocentric approaches. The virtual hand enables an isomorphic relationship between the actual and virtual hands [21]. Several works have utilized the virtual hand approach to interact with virtual elements [2], [22], [23]. The virtual pointer allows the selection of target items that are outside the user's reach with less actual hand movement. Although the virtual pointer consists of two groups [24], we will only discuss related works where the selection ray starts at the user's hand (e.g. raycasting and Go-Go [25]).

Raycasting permits selection at a distance, although the object appears smaller due to the increase in the distance of the virtual object toward the user. The virtual targeted object is selected when collision between the ray and the bounding box of the 3D objects in the environment is detected. Raycasting implementations differ in many ways, including the manner in which the ray is managed. A point of origin and a direction are needed to control the ray [26].

Table 1. Comparison of recent approaches from related research on target selection

Year	Researchers	Interaction metaphor proposed for target selection	Domain	Target object	
				Occluded	Distant
2022	This study	Gesture-based pointing with raycasting metaphor	Handheld AR	✓	✓
2022	Kapinus *et al*. [12]	Touch-based pointing with raycasting metaphor	Handheld AR	✓	-
2021	Li *et al*. [9]	Controller-based pointing with mirror metaphor	VR	✓	✓
2021	Messaci *et al*. [10]	Gesture-based with zoom metaphor	VR	✓	✓
2020	Qian *et al.* [27]	Gesture-based metaphor	Handheld AR	-	✓
2020	Yu *et al*. [28]	Controller-based pointing metaphors	VR	✓	-
2020	Sidenmark *et al*. [29]	Controller-based pointing with gaze-assisted metaphor	VR	✓	-
2019	Yin *et al*. [30]	Touch-based pointing metaphors	Handheld AR	✓	-
2018	Whitlock *et al*. [8]	Gesture-based metaphor	AR	✓	-
2017	Jung and Woo [13]	Raycasting with target object duplication metaphor	AR	-	✓
2017	Bellarbi *et al*. [14]	Gesture-based pointing with zooming metaphor	AR	-	✓
2016	Yu and Kim [31]	Finger-pointing metaphor	AR	-	✓

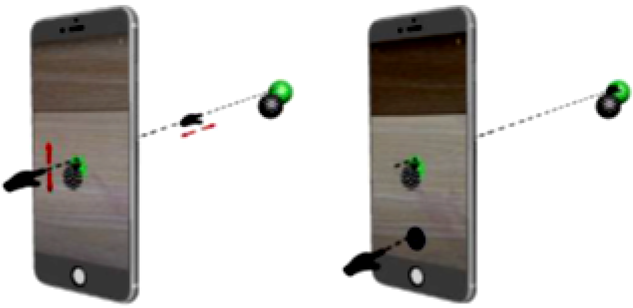
In the study by Pouprev *et al*. [25], to highlight the issues of target selection on the distant object, they prosent in their study a technique, named Go-Go which allows for seamless direct manipulation of both near and distant object. Moreover, Go-Go enables the extension of the user's ability to reach faraway object for interaction by extending a ray with the virtual hand towards the object, however, it is implemented within VR environment. Several enhancements have been made to it by other researchers throughout the years. Jung and Woo [13] have enhanced the Go-Go interaction system, by enabling it to switch on and off with trigger motions for interaction within AR environment setup, between regular and boost modes using hand gesture. Figure 1 shows that the ray is projected into a distance, where the virtual hand avatar works just like a real hand, allowing the selection of a distant virtual item, as illustrated in Figure 1(a).



(a) (b)

Figure 1. Projection of ray into the distance: (a) with virtual hand [13] and (b) with pointing raycast [32]

Meanwhile, Yin *et al*. [30] have included Go-Go as the baseline technique to be compared with their proposed novels target selection techniques (refer to Figure 2). As the interaction takes place in a handheld AR interface, a virtual hand is assigned to point toward the desired target with a single touch on the screen, conduct swipe up and down movements to alter the arm reach (see Figure 2(a)), and press the confirmation button on the screen to choose the target that has intersected with the virtual hand, as shown in Figure 2(b). However, because it limits the intended target size with a point cursor, using the virtual hand to point at and touch the obstructed or tiny target is challenging.



(a) (b)

Figure 2. The selection process for Go-Go consists of two subtasks: (a) indicating and (b) the optional step of confirming the selection [30]

Meanwhile, Olwal and Feiner [33], have presented a flexible pointer to address the occlusion issue for target selection. The target selection is done by pointing a ray cursor that can be bent to point to the user's desired target without passing through distractor targets. This method, however, requires the employment of two 6-degree of freedom devices to control the cursor, as well as the user specifying the 3D position of the intended target. Other than that, Hincapié-Ramos *et al*. [34] implemented raycasting (originating from the user's chin) based on the corresponding rotation values obtained from the handheld controller's inertial measurement unit (IMU) in an AR HMD display setup named as GyroWand. In certain implementations, the selection is activated as the ray intersects with the targeted object (e.g., Yusof *et al*. [32]) as shown in Figure 1(b). In other cases, the selection is activated by a secondary "commit" motion, such as hovering it over the target or clicking a button on a different controller [30]. Additionally, raycasting often needs a mechanism for disambiguation across multiple possible targets, especially in densely populated virtual environments. Another example is the first-generation Microsoft HoloLens device that employs a variant of the raycasting technique, in which a ray is cast from the centre of the device's viewport (accomplished by moving the head).

In addition, raycasting often necessitates a disambiguation mechanism among multiple possible targets, particularly in densely populated virtual environments. A virtual ray is cast from the point of origin, normally a tracked hand or handheld joystick, along a specified pointing direction in the default implementation. The most common approach is to implement it with a button to validate the selection. Another important distinction is the ray's form, which may either be a straight line in 3D space or have an aperture angle that basically configures a cone [21]. From Table 1, it can be seen that the approach to highlight the target selection issues are often discussed in the VR environment. Other than that, only a few researches are done to highlight the target selection issues on the virtual object with both conditions of occluded and distant, while none of the studies addressed both issues that were highlighted in handheld AR environment within their study. Therefore, this research aim is to propose a target selection method on the distant and occluded object in the handheld AR environment as another valuable contribution to the research community.

3. METHOD

The proposed target selection method is designed to address selection issues on the distant and occluded target object in handheld AR. In the proposed target selection method, the raycasting technique and hand gesture are utilized. There are several phases outlined to achieve the aim of this study which will be discussed in the next subsection.

3.1. Handheld AR workspace setup

In this study, a markerless tracking technique. SLAM technique is implemented to enable the AR environment in the handheld setup. The handheld AR workspace setup for this study is shown in Figure 3, where the user's left hand is holding the handheld device while the other hand is available for the interaction. The handheld device is required to have a camera which will be used to view the environment and gyroscope function for the tracking purpose. We implement ARCore SLAM in which the visual information (feature points SLAM captured by the camera) is used in conjunction with inertial data from the device's IMU to determine the camera's location and orientation in relation to the world over time. It searches for feature point clusters that appeared to be on identical horizontal surfaces, such as tabletops and desks, to model both horizontal and vertical surfaces represented as 3D planes. It also provides the boundaries of the plane by detecting the edges in the camera images. By aligning the pose of the virtual camera that renders the 3D content with the pose of the device's camera in the system, the virtual content is able to be rendered from the correct perspective. The more precisely the cameras are superimposed, the more credibly and more realistic the placement of the virtual object in the environment. The produced virtual images can be superimposed on top of the images collected by the device's camera, making the virtual content appear as a part of the actual environment [35].

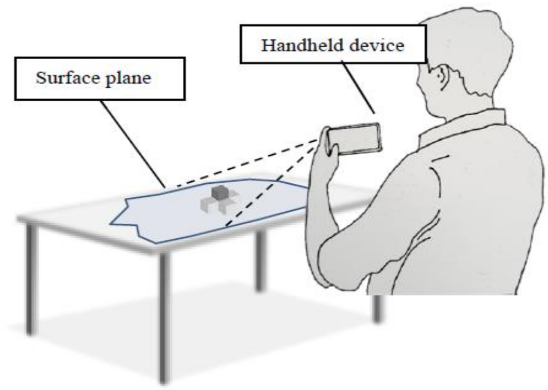


Figure 3. Handheld AR workspace setup

3.2. Enabling hand gesture tracking

The leap motion device is implemented for its advantage of producing robust hand tracking [36]. Rather than setting the leap motion sensor on the table, for this study, it is attached to the back of the handheld device, enabling intuitive and natural AR experiences. Leap motion is able to track both hands simultaneously.

However, in this study, only one hand is tracked for the interaction which is the right hand. The gesture used is the dynamic gesture, as it contains movement and motion while performing the target selection. As the leap motion projects the 3D hand gestures in different coordinates compared to the ARCamera coordinates, it needs to undergo a calibration mapping process which we adapt from Kim and Lee [2].

Figure 4 shows the depth threshold and motion-based process as the hand is recognized. This standard tracking process produced the position and orientation to form a coordinate system as the hand inputs can be used to construct the interaction methods. As the user moves their hand, leap motion can continuously track the movement. The leap motion device processes these raw data to determine the hand's skeletal calibration, which then is processed for skeletal hand tracking. Based on the raw data collected from the sensor device, it can then be used to develop features that allow the recognition of hand gestures.

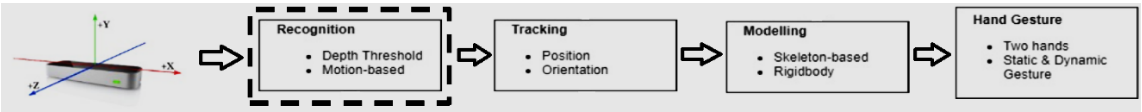


Figure 4. Leap motion–the depth threshold and motion-based process

3.3. Transferring gesture data in handheld

The leap motion device cannot be directly connected to the handheld device due to the device's limitations, as it is not designed for handheld purpose. Therefore, the gesture data that leap motion has captured needs to be sent over the network to the AR scene. The handheld device needs to ensure internet performance is in a stable bandwidth to make gesture movement smooth and robust. Otherwise, the data transmission can be delayed to obtain real-time hand movement if the network is weak. To accomplish this, a multiplayer networking is enabled in the system, adopted from the study by Nor'a *et al*. [37]. In this study, photon unity networking (PUN) is utilized to enable data transmission between devices where the attached leap motion device is connected to the Laptop (computer) to recognize the hand gesture and send the fingers' position and orientation to the server. The system running on the handheld device receives the data from the server and handles the interactions. Figure 5 shows the flow of the data transfer process.

The 3D hand orientation, gesture, and direction were obtained from leap motion which is connected to the computer. The system running on the computer acted as a Sender that transmit the position of fingertips to the PUN dedicated server. The system which was running on the handheld device, which acts as the client, then receives the hand data through the network from the server. This real-time synchronization via PUN enables the hand data from the leap motion device can be smoothly viewed on the client side through internet connectivity. Furthermore, it enables the user to further interact with the target object viewed through the handheld display using a hand gesture of the right hand.

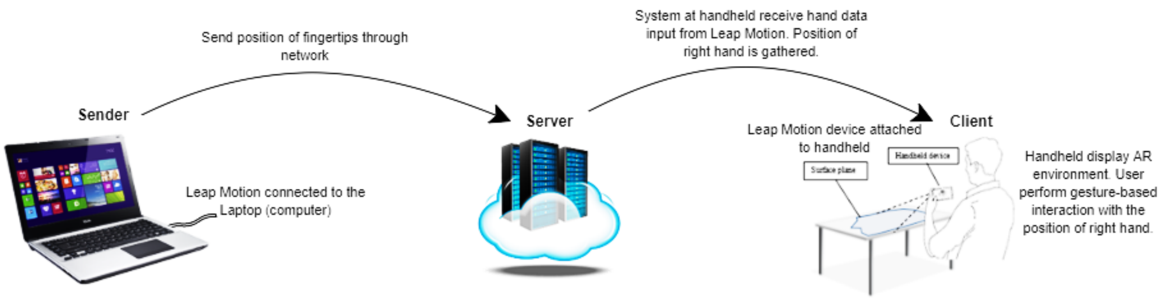


Figure 5. Gesture data transferred by sender and received at client

3.4. Proposed target selection method

After we study the related researches, the raycasting technique is proposed to be implemented with the real hand gesture for the target selection method in this study. The virtual object is selected when the collision between the ray and the bounding box of the 3D objects in the environment is detected, and the target distance range, and device tilt angle is met. Within the target selection method, it consists of two subtasks which are i) indicating the target object and ii) confirming the target selection.

The first step to performing target selection on the virtual object is initiated when the ray projected, collides with the collider of the targeted object. The virtual targeted object is selected when the collision between the ray and the bounding box of the 3D objects in the environment is detected. In this study, the ray is not simply cast directly from a position on the user's hand, instead, the handheld device position was utilized as a reference position, added an offset to approximate a shoulder position, and then projected a ray from the shoulder through the palm position and out towards the targeted object as shown in Figure 6.

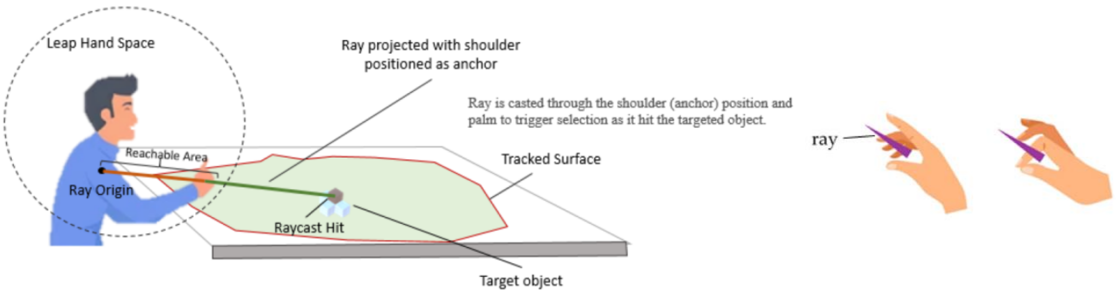
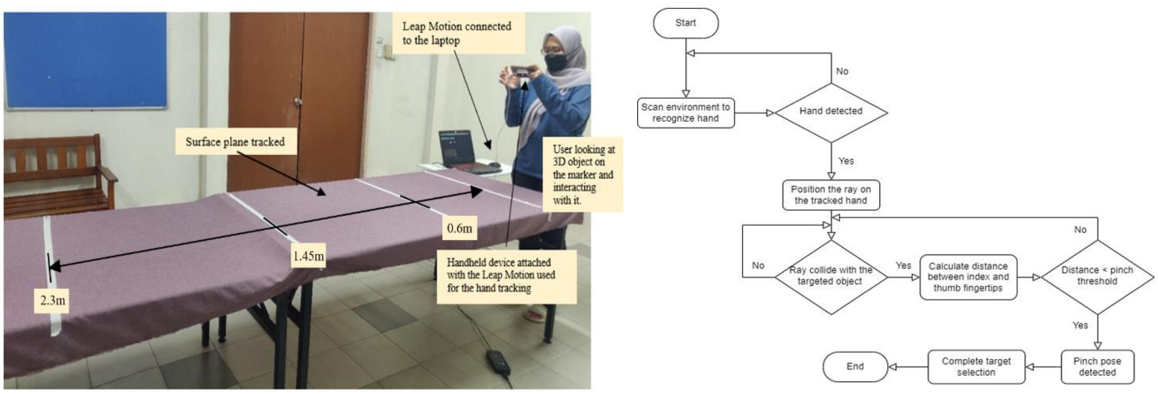


Figure 6. (Left side) illustration of the ray projected from offset shoulder as anchor, (right side) pinch gesture

The last step, which is to confirm the target selection is done with pinch gesture (refer to Figure 6 (right side)). Pinch gestures are by far the most natural ways to interact with digital interfaces. It is equivalent to grabbing or picking and provides a natural cue for selecting or moving an item in an interactive system, ensuring accuracy and high efficiency. As the leap motion device detected the hand movement, the distance between the tip of the index finger and thumb finger is calculated. If the distance between the two is less than the pinch threshold value, the pinching state of the hand gesture is detected by the system [38]. This pinching state does not change anything if there is no target object selected (the ray is not intersecting any targeted object).

4. RESULTS

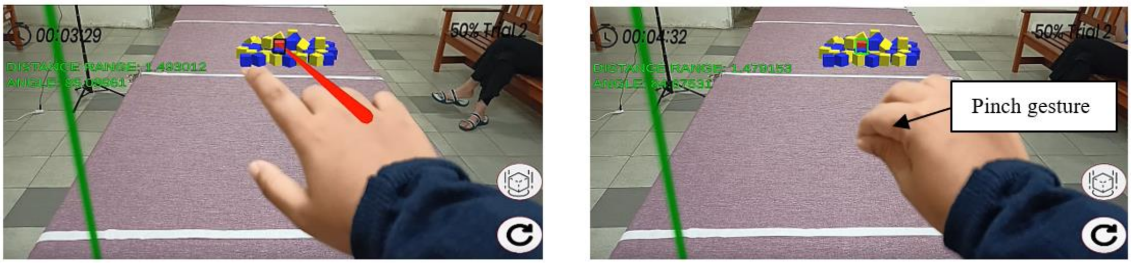
Although interaction with the virtual object includes selection and manipulation, in this study, we focus on target selection for the distant and occluded object in handheld AR. Figure 7 describes the experiment for the target selection method. As shown in Figure 7(a), the target selection is performed at three distance ranges, which are 0.6 meters, 1.45 meters, and 2.3 meters. The distance is between the centre of the handheld device and the occluded target object. The study by Qian *et al*. [27] was used as a reference as interaction performed on smartphones were studied at two interaction depths (close-range and distant), which are very close to this study. Other than that, the target object is set to be occluded by 20%, 50% and 80% of occlusion levels as referred to the study by Yin *et al*. [30]. The occluded area of the targeted object is calculated with the formula of the surface area of a cube. We further limit the range of tilt angle of the device to be within 45° to 90° for the target selection. We record the task completion time of the target selection task, to measure the performance of the target selection method. Figure 7(b) shows the flowchart of the target selection process.



(a) (b)

Figure 7. Target selection experiment (a) workspace setup and (b) flowchart of the target selection method

The right hand is utilized to perform the target selection. The process starts with scanning the environment to find the real hand. Once the real hand is detected, the condition for target selection is to check whether the ray has collided with the target object. However, before the condition is checked, once the hand is detected, a ray is drawn and positioned at the user's palm. Figure 8 shows the output of the target selection method. If the ray has collided with the target object, the target object is indicated for selection as shown in Figure 8 (a). Next, to confirm the selection, the distance between the fingertip of the thumb and index finger is calculated. Thus, the condition of the distance between thumb and index fingertips is less than the pinch threshold, which *0.4f* is checked. The red cube is the occluded target object for selection, while the blue and yellow cube is the disruptor that occluded the target object in the scene. If a target object is indicated, the object initiates visual feedback (black coloured outline of the object and once the condition has been met, and pinch gesture is detected, visual feedback is given to the user (outline of the targeted object appears in green) as shown in Figure 8(a) and (b). It is equivalent to grabbing or picking and provides a natural cue for selecting or moving an item, ensuring accuracy and high efficiency.



(a) (b)

Figure 8. Subtasks of the target selection method: (a) indicating target object and (b) confirming selection

5. CONCLUSION

This paper discussed the target selection method that we proposed using hand gesture-based with raycasting technique where the AR environment is enabled using markerless tracking technique in handheld setup. This study has addressed the issue of target selection on the distant and occluded object in handheld AR. Target selection is different from basic selection as basic selection can be made to any object; meanwhile, target selection is made to a targeted object with a certain condition. In particular, the proposed target selection method can be utilized in various fields that require interaction such as games, simulations, training, and also for educational purposes. The proposed target selection method able to perform the target selection with markerless tracking technique on the distant and occluded object in handheld AR is the main contribution. The tracking to enable AR is achieved using the SLAM tracking technique; user is required to scan their surroundings for plane surfaces. The process continues by enabling the hand gesture tracking process in the system where the hand data is transmitted through internet networking to enable the target selection process. The target selection process is designed with two subtasks. Firstly, the user will indicate the targeted object using raycast, by pointing the ray towards the targeted object. To complete the target selection, the user needs to pose a pinch gesture. To test its functionality, the time taken to perform the target selection is taken.

Therefore, for future enhancement to this study, we will conduct a complete experiment to further discuss on the performance of the proposed target selection method when performed on the occluded and distant object in handheld AR. A system usability scale (SUS) questionnaire will also be conducted after the user has completed the experiment. However, in future work, there are a few limitations that could be addressed by other researchers in their study. Firstly, the target selection method proposed in this study which implements gesture-based interaction is enabled without taking the occlusion issue for the real hand gesture into consideration. Thus, while performing selection, the virtual object augmented in the handheld AR environment always appears above the real hand gesture although the actual position for the virtual objects is behind the real hand gesture position. When performing interaction with real hand gesture recognition, the occlusion issue is an inevitable problem that needs improvement. Secondly, further study of the target selection method that is established for collaborative setup between two or more users would potentially help to support complex collaboration for professional use when interacting with a virtual object in AR. Other than that, it is suggested for future researchers to explore the adoption of different types of gestures or add speech for target selection purposes as in this study, the pinch gesture is enabled for the target selection method.

ACKNOWLEDGEMENTS

Special appreciation to Mixed and Virtual Reality Laboratory (mivielab), ViCubeLab at Universiti Teknologi Malaysia (UTM) for the adequate support and provided facilities. This work has been supported by the Ministry of Higher Education under FRGS, Registration Proposal No: FRGS/1/2020/ICT10/UTM/02/1.

REFERENCES

[1] J. M. Mota, I. Ruiz-Rube, J. M. Dodero, and I. Arnedillo-Sánchez, "Augmented reality mobile app development for all," *Computers and Electrical Engineering*, vol. 65, 2018, doi: 10.1016/j.compeleceng.2017.08.025.

[2] M. Kim and J. Y. Lee, "Touch and hand gesture-based interactions for directly manipulating 3D virtual objects in mobile augmented reality," *Multimed Tools Appl*, vol. 75, no. 23, pp. 16529–16550, 2016, doi: 10.1007/s11042-016-3355-9.

[3] V. Fuvattanasilp *et al.*, "SlidAR+: Gravity-aware 3D object manipulation for handheld augmented reality," *Computers and Graphics (Pergamon)*, vol. 95, 2021, doi: 10.1016/j.cag.2021.01.005.

[4] N. A. B. A. Halim and A. W. B. Ismail, "ARHome: Object selection and manipulation using raycasting technique with 3D-model tracking in handheld augmented reality," in *2021 IEEE 6th International Conference on Computing, Communication and Automation, ICCCA 2021*, 2021, doi: 10.1109/ICCCA52192.2021.9666416.

[5] T. Taketomi, H. Uchiyama, and S. Ikeda, "Visual SLAM algorithms: A survey from 2010 to 2016," *IPSJ Transactions on Computer Vision and Applications*, vol. 9, 2017, doi: 10.1186/s41074-017-0027-2.

[6] M. K. Bekele, R. Pierdicca, E. Frontoni, E. S. Malinverni, and J. Gain, "A survey of augmented, virtual, and mixed reality for cultural heritage," *Journal on Computing and Cultural Heritage*, vol. 11, no. 2, 2018, doi: 10.1145/3145534.

[7] D. Yu, H. N. Liang, F. Lu, V. Nanjappan, K. Papangelis, and W. Wang, "Target selection in head-mounted display virtual reality environments," *Journal of Universal Computer Science*, vol. 24, no. 9, pp. 1217–1243, 2018.

[8] M. Whitlock, E. Harnner, J. R. Brubaker, S. Kane, and D. A. Szafir, "Interacting with distant objects in augmented reality," *25th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2018 - Proceedings*, no. June, 2018, pp. 41–48, doi: 10.1109/VR.2018.8446381.

[9] N. Li, Z. Zhang, and C. Liu, "Vmirror: Enhancing the interaction with occluded or distant objects in vr with virtual mirrors," in *Conference on Human Factors in Computing Systems - Proceedings*, 2021, doi: 10.1145/3411764.3445537.

[10] A. Messaci, N. Zenati, M. Belhocine, and S. Otmane, "Zoom-fwd: Efficient technique for 3D gestual interaction with distant and occluded objects in virtual reality," *Comput Animat Virtual Worlds*, 2021, doi: 10.1002/cav.2033.

[11] N. A. B. A. Halim and A. W. B. Ismail, "Raycasting method using hand gesture for target selection on the occluded object in handheld augmented reality," in *2021 6th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, 2021, vol. 6, pp. 1–6.

[12] M. Kapinus, D. Bambušek, Z. Materna, V. Beran, and P. Smrž, "Improved indirect virtual objects selection methods for cluttered augmented reality environments on mobile devices," in *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, 2022, pp. 834–838, doi: 10.1109/HRI53351.2022.9889374.

[13] W. Jung and W. T. Woo, "Duplication based distance-free freehand virtual object manipulation," *Proceedings - 2017 International Symposium on Ubiquitous Virtual Reality, ISUVR 2017*, pp. 10–13, 2017, doi: 10.1109/ISUVR.2017.12.

[14] A. Bellarbi *et al.*, "A 3D interaction technique for selection and manipulation distant objects in augmented reality," *2017 5th International Conference on Electrical Engineering - Boumerdes, ICEE-B 2017*, 2017, pp. 1–5, doi: 10.1109/ICEE-B.2017.8192012.

[15] V. Nanjappan, R. Shi, H. N. Liang, H. Xiao, K. K. T. Lau, and K. Hasan, "Design of interactions for handheld augmented reality devices usingwearable smart textiles: Findings from a user elicitation study," *Applied Sciences (Switzerland)*, vol. 9, no. 15, 2019, doi: 10.3390/app9153177.

[16] A. U. Batmaz and W. Stuerzlinger, "Effect of fixed and infinite ray length on distal 3D pointing in virtual reality," *Conference on Human Factors in Computing Systems - Proceedings*, 2020, doi: 10.1145/3334480.3382796.

[17] M. Billinghurst, H. Kato, and I. Poupyrev, "Tangible augmented reality," *ACM SIGGRAPH ASIA 2008 Courses, SIGGRAPH Asia'08*, 2008, doi: 10.1145/1508044.1508051.

[18] A. W. Ismail and M. S. Sunar, *Multimodal Fusion: Gesture and Speech Input in Augmented Reality Environment*, vol. 331. Springer International Publishing, 2015, doi: 10.1007/978-3-319-13153-5_24.

[19] M. Y. F. Aladin and A. W. Ismail, "Designing user interaction using gesture and speech for mixed reality interface," *International Journal of Innovative Computing*, vol. 9, no. 2, pp. 71–77, 2019, doi: 10.11113/ijic.v9n2.243.

[20] I. Poupyrev, S. Weghorst, M. Billinghurst, and T. Ichikawa, "Egocentric object manipulation in virtual environments: Empirical evaluation of interaction techniques," *Computer Graphics Forum*, vol. 17, no. 3, pp. 41–52, 1998, doi: 10.1111/1467-8659.00252.

[21] F. Argelaguet and C. Andujar, "A survey of 3D object selection techniques for virtual environments," *Computers and Graphics (Pergamon)*, vol. 37, no. 3, pp. 121–136, 2013, doi: 10.1016/j.cag.2012.12.003.

[22] F. Ababsa, J. He, and J. R. Chardonnet, "Combining hololens and leap-motion for free hand-based 3d interaction in mr environments," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12242 LNCS, pp. 315–327, 2020, doi: 10.1007/978-3-030-58465-8_24.

[23] R. Radkowski and C. Stritzke, "Interactive hand gesture-based assembly for augmented reality applications," *ACHI 2012 - 5th International Conference on Advances in Computer-Human Interactions*, 2012, pp. 303–308.

[24] F. Argelaguet and C. Andujar, "Efficient 3D pointing selection in cluttered virtual environments," *IEEE Comput Graph Appl*, 2009, doi: 10.1109/MCG.2009.81.

[25] I. Poupyrev, T. Ichikawa, M. Billinghurst, and S. Weghorst, "The Go-Go interaction technique: Non-linear mapping for direct manipulation in VR," *In Proceedings of the 9th annual ACM symposium on User interface software and technology*, 1996, pp. 79–80.

[26] K. Özacar, J. D. Hincapié-Ramos, K. Takashima, and Y. Kitamura, "3D selection techniques for mobile augmented reality head-mounted displays," *Interact Comput*, vol. 29, no. 4, pp. 579–591, 2017, doi: 10.1093/iwc/iww035.

[27] J. Qian, D. A. Shamma, D. Avrahami, and J. Biehl, "Modality and depth in touchless smartphone augmented reality interactions," *IMX 2020 - Proceedings of the 2020 ACM International Conference on Interactive Media Experiences*, 2020, doi: 10.1145/3391614.3393648.

[28] D. Yu, Q. Zhou, J. Newn, T. Dingler, E. Velloso, and J. Goncalves, "Fully-occluded target selection in virtual reality," *IEEE Trans Vis Comput Graph*, pp. 1–1, 2020, doi: 10.1109/tvcg.2020.3023606.

[29] L. Sidenmark, C. Clarke, X. Zhang, J. Phu, and H. Gellersen, "Outline pursuits: Gaze-assisted selection of occluded objects in virtual reality," *Conference on Human Factors in Computing Systems - Proceedings*, 2020, pp. 1–13, doi: 10.1145/3313831.3376438.

[30] J. Yin, C. Fu, X. Zhang, and T. Liu, "Precise target selection techniques in handheld augmented reality interfaces," *IEEE Access*, vol. 7, pp. 17663–17674, 2019, doi: 10.1109/ACCESS.2019.2895219.

[31] J. E. Yu and G. J. Kim, "Blurry (Sticky) finger: Proprioceptive pointing and selection of distant objects for optical see-through based augmented reality," *International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments, ICAT-EGVE 2016*, 2016, pp. 49–56, doi: 10.2312/egve.20161434.

[32] C. S. Yusof, N. A. A. Halim, M. N. A. Nor'A, and A. W. Ismail, "Finger-ray interaction using real hand in handheld augmented reality interface," *IOP Conf Ser Mater Sci Eng*, vol. 979, no. 1, 2020, doi: 10.1088/1757-899X/979/1/012009.

[33] A. Olwal and S. Feiner, "The flexible pointer: An interaction technique for selection in augmented and virtual reality," *Uist'03*, vol. 03, pp. 81–82, 2003

[34] J. D. Hincapié-Ramos, K. Özacar, P. P. Irani, and Y. Kitamura, "Gyro wand: IMU-based raycasting for augmented reality head-mounted displays," *Proceedings of the 3rd ACM Symposium on Spatial User Interaction*, 2015, pp. 89–98, doi: 10.1145/2788940.2788947.

[35] B. Ivanova and T. Vassilev, "A mixed reality approach to visualizing cultural heritage artefacts: Mixed reality approach to cultural heritage," in *ACM International Conference Proceeding Series*, 2021, doi: 10.1145/3472410.3472432.

[36] A. Vysocký *et al.*, "Analysis of precision and stability of hand tracking with leap motion sensor," *Sensors (Switzerland)*, 2020, doi: 10.3390/s20154088.

[37] M. N. A. Nor'a, F. E. Fadzli, A. W. Ismail, Z. S. O. Vicubelab, M. Y. F. Aladin, and W. A. A. W. Hanif, "Fingertips Interaction Method in Handheld Augmented Reality for 3D Manipulation," *2020 IEEE 5th International Conference on Computing Communication and Automation, ICCCA 2020*, 2020, pp. 161–166, doi: 10.1109/ICCCA49541.2020.9250913.

[38] J. Botev, J. Mayer, and S. Rothkugel, "Immersive mixed reality object interaction for collaborative context-aware mobile training and exploration," *Proceedings of the 11th ACM Workshop on Immersive Mixed and Virtual Environment Systems, MMVE 2019*, 2019, pp. 4–9, doi: 10.1145/3304113.3326117.

BIOGRAPHIES OF AUTHORS

Ajune Wanis Ismail is a senior lecturer at Universiti Teknologi Malaysia (UTM), Johor Malaysia. She received her B.Sc., M.Sc., and Dr. Sc. degrees from UTM. She completed her M.Sc. research mode degree in a short period of 3 semesters. In 2013, she joined the Human Interface Technology Laboratory New Zealand (HITLabNZ) at the University of Canterbury as a researcher and completed her research for her PhD in 3 years. She obtained her Ph.D. from UTM in 2016, and she is now acting as the head of mivielab (Mixed and Virtual Environment Research Lab) at UTM. She keeps continuing her interest in augmented reality and mixed reality environments for her current research interest. She can be contacted at email: ajune@utm.my.

Nur Ameerah Abdul Halim is currently pursuing her master (M.Sc) at Universiti Teknologi Malaysia (UTM), Johor Malaysia under the supervision of Dr Ajune Wanis Ismail. Her research interest includes computer graphics and computer vision. She received her degree in Computer Science (Graphics and Multimedia Software) from UTM in 2015. She can be contacted at email: new.ameerah@gmail.com.

Rohaya Talib is a senior lecturer in the Department of Educational Foundation and Social Science, Faculty of Education, Universiti Teknologi Malaysia, Johor. A competent academician with more than 21 years of experience in both teaching and research. Her areas of competence and interest include educational measurement and assessment, scaling and instrumentation, quantitative statistical analysis, supervision in education, and research methodology. She is active in various research initiatives and educational breakthroughs. She is now the coordinator of the Faculty of Education's measuring and evaluation program. She can be contacted at email: rohayatalib@utm.my.

Ahmad Johari Sihes is a senior lecturer in the Department of Educational Foundation and Social Science, Faculty of Education, Universiti Teknologi Malaysia, Johor. He graduated with a bachelor's degree in chemistry from the University of Minnesota Duluth in 1986 and complete his master's degree in Curriculum Construction in 1988. He received his Ph.D. in Curriculum and Instruction from Universiti Teknologi Malaysia in 2009. He has earned various awards and distinctions, including the Science Education Award, the Best Paper in the National Student Development Conference in 2013, and the Excellence Service Award in 2016. He can be contacted at email: p-joha@utm.my.