International Journal of Integrated Engineering:
Special Issue 2018: Data Information Engineering, Vol. 10 No. 6 (2018) p. 159-168.
© Penerbit UTHM DOI: https://doi.org/10.30880/ijie.2018.10.06.023

An Inertial Device-based User Interaction with Occlusion-free Object Handling in a Handheld Augmented Reality

Goh Eg Su[1,2,3*], Mohd Sharizal Sunar[1,2], Rino Andias[4], Ajune Wanis Ismail[1,2]

[1]UTM-IRDA Digital Media Centre, Media and Game Innovation Centre of Excellence
Institute of Human Centered Engineering, Universiti Teknologi Malaysia.
[2]Department of Software Engineering, School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia.
[3]Kolej Komuniti Tanjung Piai
45-47, Jalan Delima, Pusat Perdagangan Pontian, Pontian, Johor.
[4]School of Industrial Engineering, Telkom University,
40257 Bandung, West Java, Indonesia.

Received 28 June 2018; accepted 5 August 2018, available online 24 August 2018

Abstract: Augmented Reality (AR) is a technology used to merge virtual objects with real environments in real-time. In AR, the interaction which occurs between the end-user and the AR system has always been the frequently discussed topic. In addition, handheld AR is a new approach in which it delivers enriched 3D virtual objects when a user looks through the device's video camera. One of the most accepted handheld devices nowadays is the smartphones which are equipped with powerful processors and cameras for capturing still images and video with a range of sensors capable of tracking location, orientation and motion of the user. These modern smartphones offer a sophisticated platform for implementing handheld AR applications. However, handheld display provides interface with the interaction metaphors which are developed with head-mounted display attached along and it might restrict with hardware which is inappropriate for handheld. Therefore, this paper will discuss a proposed real-time inertial device-based interaction technique for 3D object manipulation. It also explains the methods used such for selection, holding, translation and rotation. It aims to improve the limitation in 3D object manipulation when a user can hold the device with both hands without requiring the need to stretch out one hand to manipulate the 3D object. This paper will also recap of previous works in the field of AR and handheld AR. Finally, the paper provides the experimental results to offer new metaphors to manipulate the 3D objects using handheld devices.

Keywords: 3D object manipulation, interaction, device movement, handheld AR

1. Introduction

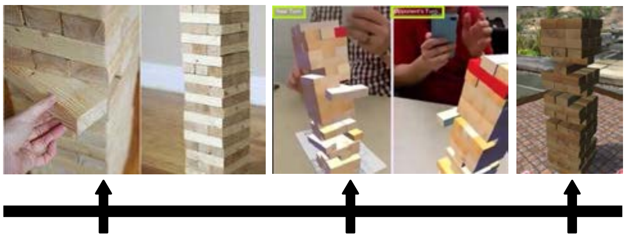
In the recent years, smartphones which are a type of handheld device is becoming a necessity in the daily life of human as it has become an essential communication tool. In addition, the incorporation of advanced systems such as the high-resolution camera equipped with different types of sensors and the installation of more powerful processors [1], makes the handheld AR potential to be further explored. Besides, the increasing demand for handheld devices makes it practical to enable the AR technology to the handheld users [2]. However, several issues related to handheld AR should be given due consideration in order to enhance the AR experience. Issues in the 3D object manipulation for user interaction still needs to be improved to enable AR in a handheld platform where most of the smartphones and tablets are limited to small screen resolution to display AR. The drawbacks in handheld displays are that it only allows 2D pointing and tapping on touch screen [3]. Meanwhile, the other problem is when the user needs to hold the device with one hand and stretching out the other to manipulate the virtual, causing fatigue upon holding the device for long [4] besides it is a lack of intuitiveness.

In handheld AR, applications are often limited to pure 2D pointing and clicking on the devices' touch-screen perspective such as the use of 2D touchscreen input as well as a stylus pen, keypad, keyboard and device sensors to interact with the virtual objects [5]. However, interaction through these 2D surfaces encounters several limitations such as screen occlusion, limited screen size and using 2D input for 3D interaction which goes against the aim of AR systems. Hence, a new user interface is needed to provide a natural, intuitive and seamless interaction experiences for the users.

*Corresponding author: egsu_goh@hotmail.com

In order to increase the intuitiveness when manipulating a virtual object, 3D gesture-based methods have been introduced [3]. By using 3D gesture tracking method, users can naturally interact with the virtual contents in 3D. In one of the earliest studies conducted in 3D gesture interaction, Henrysson et al. [6] used a fiducial marker attached to the index fingertip for tracking on the mobile phone to control a 3D painting application. In 3D interactions, users hold the handheld devices with one hand and use the other to handle the 3D object using an AR marker or the users' hands or fingers represented as AR marker to manipulate the virtual object directly within the respective camera's field of view. However, several drawbacks of this method consist of collision and incorrect virtual object's translation mismatch problem. Therefore, the issue of a handheld device held with one hand still remains unsolved. In this paper, we intend to explore the potentiality of device-based interaction method implemented by Tanikawa et al. [7], Mossel et al. [8] dan Marzo et al. [9] to improve the one-handed interaction in handheld AR.

In the current study, Jenga is proposed as a game for the application. Whereby, we demonstrate the device-based interface with the application of proposed real-time inertial reference frame method to play the simple 3D Jenga game. Jenga is a classic game, a well-known tower building/stacking game as shown in Fig. 1. This game requires precise object manipulation and a very careful interaction to avoid it from falling apart. During the game flow process, the players need to slowly pull out a block from the stack of wooden blocks tower. The player then places and balances the block on top of the stack. So in an AR simulation, inertial interaction is required to force the block to be pulled out from the stack. The inertial physics is needed to retain the balance of the tower built by the block since the user needs to control the degree of force and intensity when moving the single block from the entire tower. A slight collision with the corresponding blocks will collapse the tower.



Real Environment AR Environment Virtual Environment

Fig. 1: Jenga classic game, within its transitional from real to virtual world.

This paper will recap the previous works from relevant literature in the field of both AR and handheld AR. Based on previous literature, some benefits of handheld AR does not relate to existing handheld devices. However, these studies also focused on the interactions of inertial physics invoked to track the depth information. Handheld accelerometers are known as inertial sensors as they are allowed to exploit the property of inertia, for example, the resistance to a change in momentum in AR environment, to sense angular motion in the case of depth camera sensing and mobile changes in linear motion [10]. Furthermore, inclinometers are also inertial sensors that measure the orientation of the acceleration vector due to gravity [11]. Besides, this paper defines inertial sensors as independent of any external references or infrastructure, apart from the ubiquitous gravity field. The taxonomy presented in this paper allows a clear demonstration of where the research is focused and enables communication between researchers, designers and developers working with handheld AR technology. An interaction metaphor is proposed in this paper where handheld devices are applied as interactive tools to implement an inertial device-based tracking approach for 3D manipulation. Finally, Jenga-AR (a simple game) was developed to study the limitations of our proposed approach for 3D object manipulation.

2. Related Works

According to the virtuality continuum introduced by Milgram et al. [12], the existence of augmented reality (AR) is one of the main elements apart from real environment, augmented virtuality and virtual environment. This definition has been extended by Azuma [13] where AR technology should include not only visual input but also the other sensory inputs such as sound, taste, touch or smell. AR is known as an emerging technology to provide a potential solution for natural interaction that allow users to seamlessly interact with virtual contents that been overlaid with our real world in a real-time [14]. AR has been defined as a system that: 1) combines the real and virtual world; 2) is interactive in real time; and 3) is registered in 3D [15] [16].

Based on the available definition and technology, Feiner et al. [17] presented the Touring Machine, the first mobile augmented reality system (MARS) (see Fig. 2 (a)). The system uses a see-through head-mounted display (HMD) with integral orientation tracker, a backpack (holding a computer, differential GPS, and digital radio for wireless web access) and a handheld computer with stylus and touchpad interface. On the other hand, Höllerer et al. [18] developed a mobile AR system that allows the user to explore hypermedia news stories that are located at the places to which they refer and to receive a guided campus tour that overlays models of earlier buildings. This was the first mobile AR system to use GPS and an inertial-magnetic orientation tracker. Thomas et al. [19] presented ARQuake, an extension to the popular desktop game Quake. ARQuake is a first-person perspective application which is based on a six degree of freedom (6DOF) tracking system using GPS (as seen in Fig. 2 (c)), a digital compass and vision-based tracking of fiducial markers. Users are equipped with a wearable computer system in a backpack, an HMD and a simple two-button input device. The game can be played indoors or outdoors where the usual keyboard and mouse commands for movement and actions are performed by movements of the user in the real environment using the simple input interface.


(a) (b) (c)

Fig. 2: (a) MARS system by Feiner et al. [17], (b) user interface by Hollerer et al. [18] and (c) ARQuake by Thomas et al. [19].

In recent years, most of the user interfaces have converged. AR presents ubiquitous emerging concepts which incorporates aspects from all of the examples explained above. Ubiquitous AR (UAR) systems are typically multi-device (requiring multimodal input and multimedia output), multi-user and distributed [20]. In mobile AR, alternatives to touch-screen based interaction take advantage of the multimodal information delivered by the sensors integrated into modern smartphones [3]. The current features in mobile devices are more advanced, where most of it comes with a camera, accelerometer and compass. In particular, the accelerometer combined with a compass can be used to get the orientation of the phone. In fact, these sensors are used to create AR by specifying where and how the virtual 3D objects are registered in a scene of the real world.

Handheld AR is a highly potential platform which can be used to implement the AR technology instead of traditional desktop or tabletop AR and the HMD-based AR. There are several methods which can be used in supporting natural interaction for 3D object manipulation in handheld AR. The interaction based on the real object was introduced to manipulate virtual objects with real objects captured by AR [21] [22]. For example, in BragFish [23], this technique was applied for a collaborative AR game, where the physical markers were used in a shared physical space for the players to communicate and interact with each other. This method is convenient for the user to operate the virtual objects by handling the corresponding real objects using his/her hand. However, retention of the same height and posture of the virtual object as the real object still constrained since those virtual objects are controlled by those real objects. As a result, this method did not provide the user with an intuitive and direct way to interact with virtual objects.

Recently, handheld devices become the main platform for most of researchers to implement 3D interaction in AR system mainly select and manipulate 3D objects. In Bai et al. [24], they found out that gestural input is being used for object manipulation (as seen in Fig. 3 (c)) instead of touch screen related interaction method while fingertip become a potential AR marker to be tracked when it is within the camera field of view. Furthermore, headphone and speaker can also be used to help the user to retrieve sound inputs and converge the waves for the interaction purpose in handheld AR. In 2012, Jung et al. [25] explored a smartphone as an AR authoring tool via multitouch based 3D interaction method that used multi-touch inputs to interact with the virtual contents in handheld AR.

Another collaborative AR game is Invisible Train [26], where multiple users use portable device (PDA) to control the virtual train and 3D objects. The Invisible Train is the first multi-user AR application for handheld devices. Soon, in 2009, Arth et al. presented a system for large-scale localization and subsequent 6DOF tracking on mobile phones [27]. This system uses sparse point clouds of city areas, FAST corners and SURF-like descriptors in memory-limited devices. On the other hand, Morrison et al. presented MapLens which is a mobile AR map using a magic lens over a paper map [28] (see Fig. 3 (d)). They conduct a broad user study in form of an outdoor location-based game. They concluded that AR features facilitate place-making by creating a constant need for referencing to the physical space. The trials showed that the main potential of AR maps lies within their use as a collaborative tool. Additionally, Hagbi et al. presented an approach to track the pose of the mobile device by pointing it to fiducials [29] (see Fig 3 (e)). Besides, SiteLens (see Fig. 3 (f)), a handheld mobile AR system for urban design and urban planning site visits [30] [31] creates "situated visualizations" that are related to and displayed in their environment.

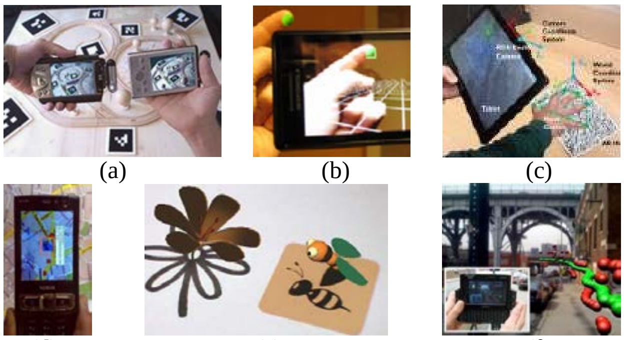

(a) (b) (c)
(d) (e) (f)

Fig. 3: (a) Wagner D. 2005 [26], (b) Hürst et al., 2013 [3], (c) Bai et al. [24], (d) MapLens, (e) Hagbi et al.[29], (f) SiteLens [30].

Due to the wide acceptation of smartphones and tablets, a hybrid method was introduced by Mossel et al. [8], which combines 3D touch interaction and device-based interaction, namely, 3D touch and HOMER-S. The latter was an extension of the research done by Henrysson et al. [32]. In these previous works, the virtual object manipulation tasks (translation and rotation) were performed separately because both tasks could not be done at the same time when a user translates the virtual object, he/she could not rotate it and vice versa. However, in HOMER-S, the user could translate and rotate virtual object integrally. According to Tanikawa et al. [7], they proposed a technique entitled the integrated view-input AR interaction for virtual object manipulation using tablets and smartphones. Their proposed approach was an extended version of Henrysson et al. [32] based on the integration of the virtual object translation and rotation tasks. Henrysson presented that the famous ARTennis game was implemented by AR technology using ARToolkit, the first collaborative AR application running on a mobile phone. For the enhancement of the users' AR experience, Tanikawa proposed a physics simulation where the users can interact more realistically. The results indicated positive reviews from users. Nevertheless, the constraints of the 3D objects rotation remain unsolved.

Table 1 describes the handheld AR timeline in the focus area of user interaction. The areas are divided into four, (1) Registration [4] [18], (2) Augmentation [8] [32], (3) User Study [3] [5] and (4) Device [7] [9]. These research areas will be detailed out in the next section with the explanation on the interaction taxonomy of handheld AR.

Table 1: The AR timeline in handheld augmented reality.

Year	Research Project	Researcher	Focus Area
2017	An Afternoon at the Museum: Through the Lens of Augmented Reality.	Schegg & Stangl [39]	Augmentation
2016	A 3D positioning Method for SLAM-based Handheld Augmented Reality.	Polvi et al. [38]	Augmentation
2015	Integrated View-Input AR Interaction for Virtual Object Manipulation using Tablets and Smartphones.	Tanikawa et al. [7]	Device
2014	Combining multi-touch input and device movement for 3D manipulations in mobile augmented reality environments.	Marzo et al. [9]	Device
2013	3DTouch and HOMER-S: intuitive manipulation techniques for one-handed.	Mossel et al. [8]	Augmentation
2012	User study of gestures in mobile AR. This study shed light on the users (U) satisfaction with registered (R) gestures.	Hürst et al. [3]	User Study
2012	Freeze view touch and finger gesture based interaction methods for handheld.	Bai, H et al. [24]	Registration
2012	Smartphone as an augmented reality authoring tool via multi-touch based 3D interaction method.	Jung et al. [25]	Device
2010	Registration (R) of user input via sketching.	Hagbi et al. [29]	Registration
2008	More realistic representation of augmentations, to study more photorealistic augmentations (A).	Nishina et al. [36]	Augmentation
2008	Camera adjusted representations of augmentations. To enhance the augmentations.	Klein et al. [37]	Registration
2007	Proposed development of device for handheld AR. A tech note on how to create an ergonomic device (D).	Kruijff et al. [35]	Device
2006	How users (U) perceive and relate to humanoid augmentations (A).	Wagner et al. [5]	User Study
2005	Virtual object manipulation using a mobile phone	Henrysson [32]	Augmentation
2003	First handheld AR, focus on registration and representations of AR	Wagner et al. [34]	Registration

On the other hand, AR interaction using handheld devices mostly studies the touchscreen design. Several researchers have explored the use of basic 3D object manipulation for multi-touch inputs [8] [9] [24] [25]. In these related works, improvement of multi-touch inputs to cover the essential 3D object manipulation tasks (holding, selecting, translating and rotating) may require more than two fingers touch [40]. These were explored to represent each DOF of 3D object translation, rotation or even scaling. Prior knowledge is needed for users to use this approach [8] [25]. Users can even scale or deform the virtual object with a two fingers touch. Direct and indirect touches are used to perform full 3D manipulation to all x, y and z-axes. Complicated manipulation can be done with the association of more than one touch at the right moment. Yet, this approach faces some limitations. Firstly, fat fingers might occlude and prevent the correct position of the 3D object being tracked. Secondly, limited touchscreen space could prevent the user from translating and rotating virtual object because of the distance (too long) and rotation (larger rotation range) [9].

3. Interaction Taxonomy of Handheld AR

The taxonomy discussed here is based on the study of related works in the previous section. It describes a clear demonstration of where the research is focused and to enable communication between researchers, designers and developers working with handheld AR technology.

3.1 Augmentation

The augmentation is crucial to give the applications its purpose. In some cases, different approaches to rendering can improve the usability of the application by adjusting the quality of the displayed image [35] or by trying to achieve greater photorealism [36]. How content and augmentation are presented and perceived [26] is a continuing effort within the field of handheld AR.

In 2016, another research was conducted based on providing meaningful augmentation in handheld AR which focuses on accurate task positioning for virtual contents in it [38]. The method introduced was named slidAR, a SLAM (simultaneous localization and mapping) based 3D positioning method utilizing 3D ray-casting [41] and epipolar geometry [42] implemented for 3D object translation and after evaluation. The method was proven to present faster task performance and higher augmentation precision. Besides, this layer borrows extensively from the current state of the art in real-time visualization of 3D graphics. By the virtue of existing handheld device, handheld AR applications inherit the current technological constraints as well as the mobile user in action. How we can create meaningful augmentations to mobile users on limited hardware in different environmental, social and cultural contexts is interesting topics for future research. A 3D model of troll figure projected in 3D represents the augmentation layer in Fig 4. However, the drawbacks of the above methods seems to have no solution for the issues in 3D object manipulation for handheld AR stated previously [15].

Nevertheless, research in the related area is still ongoing and other 3D object manipulation metaphors are currently being introduced.

As illustrated in Fig.4, the layers consist of the device, augmentation and registration that have been explicated in this section. Besides these layers, the graphic user interface (GUI) is projected between the device and augmentation. GUI is a fundamental layer in the interaction taxonomy for handheld AR. It commonly refers to the graphical icons and visual indicators such as touch button, slider and label to instruct and provide system feedbacks to the user when interacting with handheld devices. The design of GUI is important to enhance the user experience when interacting with handheld devices. On the other hand, the last layer indicates that the handheld device also has a world projection layer which refers to the practical usage of AR in the real environment. Some of the applications in handheld AR can be applied into various areas that contain real-world contents such as urban planning, heritage protection [43], medical, game and instructional tools by applying the inertial device-based interaction technique.

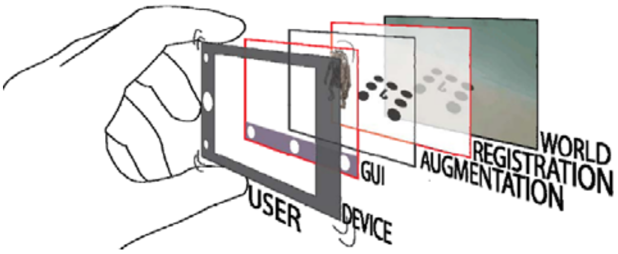


Fig. 4: Augmentation in handheld AR [8][31].

3.2 Registration

This layer is responsible for illustrating the registration of virtual contents into the real world that is an important part of a handheld AR application. The field of registration involves research which deals with the hard science of registration. This encompasses user input by the common manipulation of the marker, natural feature or any other tracked surface or object.

As depicted in Fig. 5, the virtual teapot overlays the top of the marker through a mobile where the figure explains the standard registration process in an AR tracking for handheld devices. When the tracking sensor within the built-in camera on a handheld device searches and registers the marker, the position and orientation of the potential marker will be calculated to identify the pattern. After completing the search of the list of markers that is stored in the database, the registered marker is called where the position and orientation of the marker are calculated according to the transformation matrix which corresponds with the built-in digital camera to render a 3D virtual object in the video frame. Next, the video streaming process depends on the processing time of the handheld device to display the virtual object on the visible camera view.

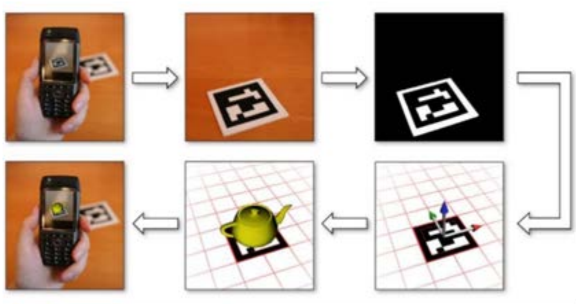


Fig. 5: Registration in handheld AR.

3.3 User Study

This layer represents the focus on users and user studies within the realm of mobile AR. With handheld being an emerging field, Swan et al. [44] showed that formal and deep user studies on AR applications are an integral part of the research contribution to the AR field in general.

According to one of the important consideration to enhance the user AR's experience is the intuitiveness, a gesture-based interaction for handheld AR [3]. By using the user's hand gestures and fingers, it is impossible to perform virtual object manipulation [3] [22] [45] [46]. This method recognizes the user's hand or fingers tracked in the camera view. By mapping the position and orientation of the user's hand and fingers with the virtual object, it can be manipulated (translated or rotated) through the movement of the user's hand and fingers. Therefore, users can naturally handle the virtual object like in the real world. As agreed by [14], the accurate tracking technique will enhance the user interaction in AR. However, the remaining issues in 3D object manipulation are lighting, occlusion problems and degree of freedom. Occlusion problems occur when the virtual object is too close to the device's camera as a message of tracking error or failed to handle the virtual object will be displayed. Detection of the user's hand and fingers become difficult when they appear in an occluded manipulation area. Moreover, this technique has been proven to lack accuracy and yield high error rates [3].

3.4 Device

Hence, from the device or hardware aspect, the current new 3D manipulation metaphors are proposed to solve some of the existing problems as a substitute for the above layers: decoupling the multi-touch interaction techniques with an interface to separate the manipulation task and to reduce the usage of multi fingers to one at a time [25]. Again, this technique is still unable to solve the limitation of touchscreen space for handheld devices [7].

Based on the proposed approach in this study, the device-based interaction aims to improve the screen space limitation and handle the difficulties when performing all 6DOF manipulations through the use of at least two finger touches for multi-touch 3D manipulation technique [8] since the position and orientation of the handheld device's built-in camera is mapped absolutely with the position and orientation of the virtual object. Besides, Tanikawa et al. [7] stated that the device-based interaction can also overcome problems with gesture-based interaction that use hands and fingers, resulting in the low accuracy in 3D object manipulation since the virtual object is moved or rotated consistently followed the handheld device's movement (mapped the position and orientation of the built-in camera). Furthermore, a user can hold the device with both hands to manipulate the virtual object since the virtual object is being manipulated (translated and rotated) when user translates and rotates the handheld device that not lead to one-hand interaction.

The remaining issues in the previous studies were the instability of hand operation when positioning the virtual object [8] [9], slow and constrained rotations of a virtual object and also the inability to create movement in the z-axis [8] [9]. On the other hand, the constraints of the device-based interactions include the slow object rotation caused by the pitch axis limitation mentioned by Mossel et al. [8] in their research and also the 360 degree z-axis (yaw) rotation that requires the user to move around, hence, slowing down the 3D object's rotation time. These limitations are also supported by Marzo et al. [9], where users lose sight of the manipulated object in the screen when using device-based method for orientation. Some other approaches use more than one 3D object manipulation techniques in handheld AR such as touch-based technique used together with gesture-based technique to improve performances. This is known as multimodal technique.

4. Proposed Method

In this paper, the inertial device-based interaction method proposed by Tanikawa et al. [7], Mossel et al. [8] and Marzo et al. [9] are explored with improvement of the virtual object selection in handheld. This proposed method uses the device's camera references frame as the inputs to provide real-time manipulation tasks in four particular interactions, consisting of selecting, holding, translation and rotation.

4.1 Jenga-AR Handheld Prototype

AR can provide users with enhanced interaction experiences by integrating virtual and real-world objects in an AR environment. Through the AR interface, a more natural and immersive control style is achievable compared to the traditional keyboard and mouse input devices. The Jenga-AR interface which is proposed in this paper consists of a stereo camera, which tracks the blocks of Jenga where the camera viewpoint accurately captures the 3D space of the blocks. Furthermore, in order to enable a physically realistic experience in the interaction, a physics engine is adapted to simulate the physics of virtual object manipulation. Traditionally, the blocks can be picked up and tossed with physical characteristics, such as gravity and collisions which occur in the real world. However, the interaction metaphor in our system is fully device-based, without markers, user's hands or fingers. Therefore, the 3D scene is overlaid into real-world using a feature-based marker.

We proposed a Jenga-AR application into the low-cost smartphone branded Huawei Y6II Cam-L21 to observe the application's performance and find out whether an additional function script is needed to enable the built-in camera to focus on the AR marker automatically. The AR interface for Jenga game was deployed in an android device with a touch button at the bottom right portion of the camera field of view of the handheld device to ease the user to hold and release the virtual object after selection (as seen in Fig. 6).

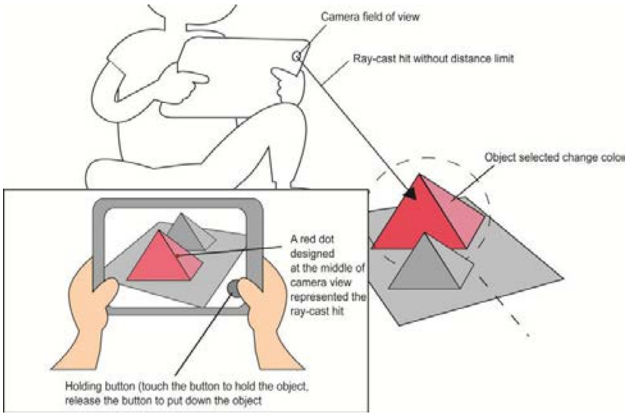


Fig. 6: Virtual object manipulation using device-based interaction method.

Before applying the interaction metaphor into Jenga-AR, the distance range in interface is initialized with zero values for the distance between the device and the 3D Jenga for selection function. By eliminating the precision of length, it would enable the interaction space for 3D object to be selected without distance limitation. The defined distance as example to represented far and near selection is illustrated in Fig. 7. Besides, since the depth (z-axis) distance between device's built-in camera position and orientation is mapped absolutely with the 3D object thus can reduce the huge calculation of depth information. This helps to reduce the usage of computing processor of the handheld device and avoids huge calculations which could exhaust the battery life.

As seen in Fig. 7, ray-casting method is used to select the block of Jenga by hitting it with traced ray. A red plot at the middle of the camera view represents the ray pointing, the virtual object that overlaid or collided with the traced ray will change the color to the red to differentiate the selected blocks with the remaining deselected blocks. User can handle the virtual object whether to select, holding, translate or rotate by holding the device with both hands. This metaphor acts naturally like holding the 3D block by hand while being able to control the 3D block that is overlaid in the real environment. However, since this method has a certain degree of limitation on the freedom of manipulation, it consists of drawbacks. The virtual objects become slow with constrained rotations when the user is required to move around. This limitation takes time and delays the interaction, where the movement in the z-axis is made impossible because the AR marker disappears quickly from the scene. Therefore, we employed several concepts applied by Tanikawa et al. [7], Mossel et al. [8] and Marzo et al. [9], for the translation and rotation of the virtual objects. Hence, a method to handle blocks was discovered. The method will be described in the next subsection.

(a) AR marker is far from camera

(b) AR marker is near from camera

Fig. 7: Device-based 3D object manipulation without distance limit.

4.2 Implementing Inertial Device-based Interaction

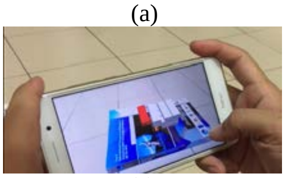
In this paper, an inertial device-based tracking method was proposed for the 3D object manipulation in handheld AR to retain the position of the interaction point relative to the device's movement. The device-based inputs are capable of performing the manipulation tasks (select, hold, translate and rotate). Here, the inertial refers to the camera view coordinate which is fixed with the selected virtual object to enable its movement within the x, y and z-axes. However, this method is improved by eliminating the ray-cast hit distance to enable the selection of virtual objects without limiting the distance calculation between the user and the virtual object. As explained in the previous sections, the Jenga-AR is designed with a touch button for the user to touch and hold the virtual object while the camera viewpoint works on to tracking the closest blocks. Jenga-AR implementation in this paper is used to test our proposed method. The method presents a device-based user interaction where the user holds the device using both hands and moves the device to track the desired virtual blocks in real-time.

The blocks are displayed on the textured image as a marker as illustrated in Fig. 8 (a). With the movement of the camera, the current viewpoint will store the reference frame after it captures the virtual blocks. The user then needs to hover over the viewpoint until it hits the block in order to perform an interaction (Fig. 8 (b)). Once, the block of Jenga is targeted by the camera tracking device and the block turns red which requires the user to pull out the block during simulation, physic is being applied to give a realistic puling experience. Then, user can moves the selected block to the top of tower by moving the handled device (as shown in Fig. 8 (c)).

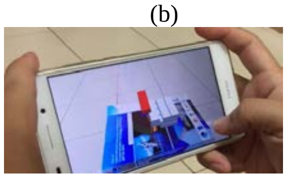
The prototype is successfully tracking the AR marker and the game still visible overlaid on the top of it even it would disappear when the AR marker not visible within the camera view. Hold button is provided to enable the user to pick up the block by touching the button on screen or releases the block by not touching it. Gravity in rigid-body modeling provides a valuable spatial reference, however for rotations about a vertical axis gravity provides no cues, and handheld AR integration is required to keep track of textured marker. Handheld device as a handling tool and its camera as a virtual tracker are running in real-time to force an inertial to the device tracking system which used to hold the virtual object. Then the user able to move the block around the environment based on the device movement. User can pick the virtual block after perform selecting and keep holding the virtual button on screen to translate by moving the device camera position (as in Fig. 8 (b, c, d)). The block will be collided with the existing blocks to complete the Jenga simulation (as in Fig. 8 (f)). When the user touches the virtual block that he wants to manipulate to avoid occlusion among user's eyes and the camera view and then ease the user to perform the manipulation task. A physic simulation system is applied to provide the user with a more realistic way when manipulate the virtual object. We design the holding button on the bottom right of the camera view to enable the user to hold the device with both hands to avoid causing fatigue after a long period of holding the device also in order to stabilize the device movement. Release the virtual button while the block correctly placed on the estimated position to complete Jenga game. Touch the virtual button and hold on the button will stay holding the picked block. User removes his thumb's finger from the screen (as seen in Fig. 9 (a)) and it will execute release action to place down the block as in Fig. 9 (f).

(a)

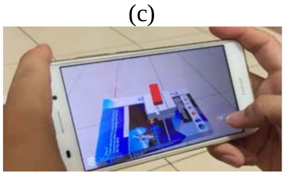
(b)



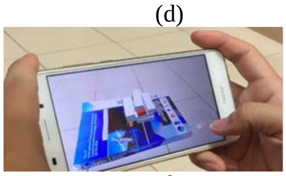
(c)



(d)

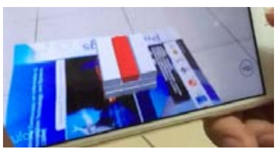


(e)



(f)

Fig. 8: Virtual object manipulation using device-based movement.



(a)

(b)

Fig. 9: Virtual object's selecting and holding tasks.

5. Conclusion

This paper discusses the Jenga-game implemented by Tanikawa et al. [7] using the device-based method. However, the developed system has not been evaluated where we plan to find a solution to improve the current method (slow 3D object rotation and constraint of the z-axis when rotating the 3D object). We have proven that the acceleration input provided by the handheld device is able to handle the 3D object, especially for the rotation task. The full potential of our proposed method can be explored to utilize its strength and usage in different areas besides 3D games.

Furthermore, some additional functions can be implemented through the 3D manipulation tasks to enable interaction cues using handheld device's pose, such as shaking the device to delete the virtual content selected. Since the handheld devices nowadays come with different kinds of sensors, it is possible to apply multimodal interactions to combine more than one input, for example, gesture and speech. Thus, it also possible to explore device-based interaction with others modality to diversify the AR implementation in different aspects and conduct a user survey on its ease of use, usability and the task completion's time.

Meanwhile, collaborative AR in multi-user works can also be a research direction. AR applications with implemented multi-user interaction as proposed by Stafford et al. [47] use full 3D capture and reconstruction to facilitate the communication of situational and navigational information between indoor users equipped with tabletop displays and outdoor users equipped with mobile AR systems. Therefore, the implementation of the inertial device-based interaction technique for a multi-user interaction especially for tasks which requires connecting different users at different locations. This will be a convenient AR interface for urban planning when involving multiple interactions in different tasks.

In conclusion, device-based interaction was discussed in detail due to its strengths and characteristics that could help in overcoming some issues faced by other methods. However, it still faces constraints in virtual object rotation tasks which needs consideration. The drawbacks of this method need to be explored in depth to find potential solutions. Device-based methods can be implemented in handheld AR where it can be useful for individuals involved in urban planning, interior design, education and furniture retail sector in the future. In this paper, we have applied the device-based method recommended by the other researchers with improvised user interface design phase to explore this method analytically and to understand its basic concepts. Nevertheless, a more sophisticated improvement is in line to overcome the speed and constraints of the virtual object rotation task. Meanwhile, other features of the device-based method can be explored to understand its practical usage in different sectors. We hope this paper could help researchers to come up with new ideas to improve or enhance the current methods and provide some useful information.

Acknowledgement

We thank the other members of the Media and Game Innovation Centre of Excellence (MaGIC-X), Universiti Teknologi Malaysia for their time, effort and enthusiasm. We express our gratitude to the Ministry of Higher Education, Malaysia (MOHE) for sponsoring and supporting us in this research where the main author was awarded a scholarship under MOHE.

References

[1] S. Kurkovsky, S., R. Koshy, V. Novak and P. Szul, "Current issues in handheld augmented reality," in: Proceedings of Communications and Information Technology (ICCIT) on IEEE International Conference (June), pp. 68-72, 2012.

[2] F. Zhou, H. Duh, M. Billinghurst, "Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR," in: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, IEEE Computer Society, pp. 193-202, 2008.

[3] W. Hürst, C.V. Wezel, "Gesture-based interaction via finger tracking for mobile aug-mented reality," Multimedia Tools and Applications 62(1), pp. 233-258, 2013.

[4] C.F. Yusof, H. Bai, M. Billinghurst, M.S. Sunar, "A review of 3D gesture interaction for handheld augmented reality," Jurnal Teknologi 78(2-2), 2016.

[5] D. Wagner, D. Schmalstieg, "Handheld augmented reality displays," in: Virtual Reality conference, 2006.

[6] A. Henrysson, J. Marshall, M. Billinghurst, "Experiments in 3D interaction for mobile phone AR," in: Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia, SER. GRAPHITE '07, ACM, pp. 187-194, 2007.

[7] T. Tanikawa,, H. Uzuka, T. Narumi, M. Hirose, "Integrated view-input AR interaction for virtual object manipulation using tablets and smart phones," in: Proceedings of the 2015 International Conference on Advances in Computer Entertainment Technology, ACM, 2015.

[8] A. Mossel, B. Venditti, H. Kaufmann, "3DTouch and HOMER-S: intuitive manipulation techniques for one-handed handheld augmented reality," in: Proceedings of the Virtual Reality International Conference: Laval Virtual, ACM, 2013.

[9] A. Marzo, B. Bossavit, M. Hachet, "Combining multi-touch input and device movement for 3D manipulations in mobile augmented reality environments," in: Proceedings of the 2nd ACM symposium on Spatial user interaction, pp. 13-16, 2014.

[10] L. Chai, W.A. Hoff, T. Vincent, "Three-dimensional motion and structure estimation using inertial sensors and computer vision for augmented reality," Presence: Teleoperators and Virtual Environments 11(5), pp. 474-492, 2002.

[11] J.M. Santana, J. Wendel, A. Trujillo, J.P. Suárez, A. Simons, A. Koch, "Multimodal location based

services—semantic 3D city data as virtual and augmented reality in progress in Location-Based Services 2016”, Springer International Publishing, pp. 329-353, 2017.

[12] P. Milgram P and F. Kishino, “A taxonomy of mixed reality visual displays,” IEICE TRANSACTIONS on Information and Systems 77(12), pp. 1321-1329, 1994.

[13] R.T. Azuma, “A survey of augmented reality. Presence: Teleoperators and virtual environments,” 6(4), pp. 355-385, 1997.

[14] A.W. Ismail, M. Billinghurst, M.S. Sunar, “Vision-based technique and issues for multimodal interaction in augmented reality,” in: Proceedings of the 8th International Symposium on Visual Information Communication and Interaction, ACM, pp. 75-82, 2015.

[15] D.W.F. van Krevelen, R. Poelman, “A survey of augmented reality technologies, applications and limitations,” International Journal of Virtual Reality 9(2), 2010.

[16] M. Mekni, A. Lemieux, “Augmented reality: Applications, challenges and future trends,” in: Proceedings of the 13th International Conference on Applied Computer and Applied Computational Science (ACACOS ‘14), pp. 205-214, 2014.

[17] S. Feiner, B. MacIntyre, T. Höllerer and A. Webster, “A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment,” in: Proceedings of First IEEE International Symposium on Wearable Computers (ISWC '97), 1997.

[18] T. Höllerer, S.Feiner, J. Pavlik, “Situated documentaries: Embedding multimedia presentations in the real world,” in: Proceedings of IEEE International Symposium on Wearable Computers, 1999.

[19] B.H. Thomas, B. Close, J. Donoghue, J. Squires, P.D. Bondi and W. Piekarski, “First person indoor/ outdoor augmented reality application: ARQuake. Personal and Ubiquitous Computing 6, pp. 75-86, 2002.

[20] O. Hilliges, “Interaction management for ubiquitous augmented reality user interfaces: Car augmented reality,” Thesis, Department of Informatics, Technische Universität München, Germany, 2004.

[21] D. Huynh, K. Raveendran, Y. Xu, K. Spreen and B. MacIntyre, “Art of defense: a collaborative handheld augmented reality board game,” in: Proceedings of the 2009 ACM SIGGRAPH symposium on video games, pp. 135-142, 2009.

[22] T. Ha, W. Woo, “An empirical evaluation of virtual hand techniques for 3D object manipulation in a tangible augmented reality environment,” in: Proceedings of 3D User Interfaces (3DUI) on 2010 IEEE Symposium, IEEE, pp. 91-98, 2010.

[23] Y. Xu, M. Gandy, S. Deen, B. Schrank, K. Spreen, M. Gorbsky, B. MacIntyre, “BragFish: Exploring physical and social interaction in co-located handheld augmented reality games,” in: Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology, pp. 276-283, 2008.

[24] H. Bai, G.A. Lee, M. Billinghurst, “Freeze view touch and finger gesture-based interaction methods for handheld augmented reality interfaces,” in: Proceedings of the 27th Conference on Image and Vision Computing New Zealand, pp. 126-131, 2012.

[25] J. Jung, J. Hong, S. Park, H.S. Yang, “Smartphone as an augmented reality authoring tool via multi-touch based 3D interaction method,” in: Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry, pp. 17-19, 2012.

[26] D. Wagner, T. Pintaric and F. Ledermann, “The invisible train; Group exhibition; organised by: Imagina, catalogue: Imagina 2005, Grimaldi Forum, Monaco, 2005.

[27] C. Arth, D. Wagner, M. Klopschitz, A. Irschandheld Ara and D. Schmalstieg, “Wide area localization on mobile phones,” in: Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR), 2009.

[28] A. Morrison, A. Oulasvirta, P. Peltonen, S. Lemmela, G. Jajucci, G. Reitmayr, J. Näsänen, A. Juustila, “Like bees around the hive: A comparative study of a mobile augmented reality map,” in: Proceedings of the Conference on Human Factors in Computing Systems, 2009.

[29] N.O. Hagbi, J. Bergig, El-Sana and M. Billinghurst, “Shape recognition and pose estimation for mobile augmented reality,” in: Proceedings of Mixed and Augmented Reality, 2009.

[30] S. White, S. Feiner, “SiteLens: Situated visualization techniques for urban site visits,” in: Proceedings of the Conference on Human Factors in Computing Systems, 2009.

[31] A.W. Ismail, and M.S. Sunar, “Collaborative augmented reality approach for multi-user interaction in urban simulation,” in: Proceedings of Information and Multimedia Technology, 2009.

[32] A. Henrysson, M. Billinghurst, M. Ollila, “Virtual object manipulation using a mobile phone,” in: Proceedings of the 2005 international conference on Augmented tele-existence, 2005.

[33] A. Henrysson, M. Billinghurst, M. Ollila, “Face to face collaborative AR on mobile phones,” in: Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2005.

[34] D. Wagner and D. Schmalstieg, “First steps towards handheld augmented reality,” in: Proceedings of the 7th International Conference on Wearable Computers, IEEE Press, pp. 127 – 135, 2003.

[35] E. Kruijff, E. Veas, “Vesp'R - Transforming handheld augmented reality,” in: Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality (SMAR), 2007.

[36] Y. Nishina, B. Okumura, M. Kanbara and N. Yokoya, “Photometric registration by adaptive high dynamic range image generation for augmented reality,” in: Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (SMAR), 2008.

[37] G. Klein and D. Murray, “Compositing for small cameras,” in: Proceedings of the 7th IEEE/ACM

International Symposium on Mixed and Augmented Reality (SMAR), 2008.
[38] J. Polvi, T. Taketomi, G. Yamamoto, A. Dey, "SlidAR: A 3D positioning method for SLAM-based handheld augmented reality," ELSEVIER: Computers & Graphics 55, pp. 33-43, 2016.
[39] R. Schegg, B. Stangl, "An Aternoon at the Museum: Though the lens of augmented reality," in: Proceedings of the International Conference in Information and Communication Technologies in Tourism, 2007.
[40] A. Martinet, G. Casiez, L. Grisoni, "Integrality and separability of multitouch interaction techniques in 3D manipulation tasks," IEEE transactions on visualization and computer graphics 18(3), pp. 369-380, 2012.
[41] B. Wang, F. Faure, D. Pai, "Adaptive image-based intersection volume," ACM Trans. Graph 31(4), pp. 1-9, 2012.
[42] A. Goldstein and R. Fattal, "Video stabilization using epipolar geometry," ACM Trans. Graph 31(5), pp. 1-10, 2012.
[43] Z. Noh, M.S. Sunar, Z. Pan, "A review on augmented reality for virtual heritage system," in: International Conference on Technologies for E-Learning and Digital Entertainment, Springer, 2009.
[44] J. Swan J and J. Gabbard, "Survey of user-based experimentation in augmented reality," in: Proceedings of 1st International Conference on Virtual Reality, 2005.
[45] H. Bai, L. Gao, J. El-Sana, M. Billinghurst, "Markerless 3D gesture-based interaction for handheld augmented reality interfaces," in: Proceedings of Mixed and Augmented Reality (ISMAR), 2013.
[46] C. Telkenaroglu and T. Capin, "Dual-finger 3D interaction techniques for mobile devices," Personal and ubiquitous computing 17(7), pp. 1551-1572, 2013.
[47] A Stafford, W. Piekarski and B. Thomas, "Implementation of god-like interaction techniques for supporting collaboration between outdoor AR and indoor tabletop users," in: Proceedings of the 5th IEEE and ACM International Symposium on Mixed & Augmented Reality, 2006.