UTM UNIVERSITI TEKNOLOGI MALAYSIA
INTERNATIONAL JOURNAL OF INNOVATIVE COMPUTING
ISSN 2180-4370
Journal Homepage : https://ijic.utm.my/

Pre-define Rotation Amplitudes Object Rotation in Handheld Augmented Reality

Goh Eg Su & Ajune Wanis Ismail
School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia
Email: eg.su@utm.my; ajune@utm.my

Submitted: 28/2/2021. Revised edition: 8/7/2021. Accepted: 11/7/2021. Published online: 15/11/2021
DOI: https://doi.org/10.11113/ijic.v11n2.315

***Abstract*—Interaction is one of the important topics to be discussed since it includes the interface where the end-user communicates with the augmented reality (AR) system. In handheld AR interface, the traditional interaction techniques are not suitable for some AR applications due to the different attributes of handheld devices that always refer to smartphones and tablets. Currently interaction techniques in handheld AR are known as touch-based technique, mid-air gesture-based technique and device-based technique that can led to a wide discussion in related research areas. However, this paper will focus to discover the device-based interaction technique because it has proven in the previous studies to be more suitable and robust in several aspects. A novel device-based 3D object rotation technique is proposed to solve the current problem in performing 3DOF rotation of 3D object. The goal is to produce a precise and faster 3D object rotation. Therefore, the determination of the rotation amplitudes per second is required before the fully implementation. This paper discusses the implementation in depth and provides a guideline for those who works in related to device-based interaction.**

***Keywords*—Rotation, Amplitudes, Pre-defined, Device-based**

I. INTRODUCTION

Currently, there are three types of interaction techniques investigated widely to be used in handheld AR which are touch-based interaction [1], [2], mid-air gestures-based interaction [3] and device-based interaction [3], [4]. Comparison has been made in several aspects and it has been proven that device-based interaction is a more suitable and robust technique [5], [6].

The selection and manipulation of 3D object that includes translating and rotating in handheld AR are the fundamental tasks in 3D interaction [1]. 3D object manipulation is commonly performed using touch-based interaction because of the widespread use of touch-screen smartphones and tablets such as IPhones, IPads, Samsung Galaxy series phones, etc. [2], [3], [4]. The drawback in handheld displays is that it only allows 2D touching on screen [7].

Moreover, the users need to have prior knowledge before performing the manipulation task using touch-based interaction technique, where different finger gestures are required to perform different manipulation tasks and usually only one finger touch is needed to do the selection and translation task while two fingers are used to perform the rotation task [8], [9]. Since the user needs to place his/her finger touch on the device screen, occlusion may occur and the user fails to view the virtual object on the display screen. Besides that, there is a lack of intuitiveness and it is not categorized as a natural interaction way [2], [3], [4].

In order to increase intuitiveness when manipulating the 3D objects, mid-air gestures-based interaction technique has been introduced [7]. By using 3D gesture tracking method, users can interact with the virtual contents in 3D naturally. The earliest studies in 3D gesture interaction, Henrysson [10] had used a fiducial marker attached to the index fingertip and tracked at the front of the mobile phone [10]. This was used to control a 3D painting application. In 3D interaction, users hold the handheld device with one hand and use another hand to handle the 3D object using an AR marker or else the hand represents the AR marker to manipulate the 3D object directly within the respective camera's field of view. However, this method contains some inferiorities such as occlusion problem and also incorrect 3D object's position deviation problem. When users hold the 3D object in mid-air, it is difficult to release it at precise position because of occlusion. Detection of the user's hand and fingers become difficult when they appear in an occluded manipulation area. Furthermore, this technique has been proven achieve high error rates [7].

Device-base interaction technique is a current interaction technique used in handheld AR that is based on video see-through AR. This technique uses information of the back-facing camera installed on the handheld device to capture the primitive data from the movement of handheld device to perform the fundamental 6 DOF virtual 3D object manipulation that includes translation and rotation tasks [3], [4], [11], [12]. Briefly, the 3D object selected will be manipulated to achieve the desired position and pose when the handler translates or rotates his/her handheld device while the handheld device acts as a holding tool or adhesion agent to pick up or stick up the 3D object on its relative part mapped with the camera field of view.

This technique is a robust technique since the 3D object is being manipulated following the device movement that considers the camera sensor input also acts as a tracking system in handheld AR therefore avoiding the occlusion problem faced by mid-air gestures-based interaction technique [11], [12].

Meanwhile, device-based interaction technique also excludes the distance limitation faced by mid-air gestures-based interaction technique and touch-based based interaction technique when translating a 3D object from one point to another while the 3D object is moved following the device movement and always seen in the camera field of view. In chief, device-based interaction technique is the only technique that allows users to hold the handheld device with both hands unlike other techniques where the device possession is limited to one hand and requires the user to stretch out one hand to manipulate the 3D object which may cause fatigue after a long period holding the handheld device single-handed [3], [4], [11], [12].

II. RELATED WORKS

Summarized from the previous section in this paper, it can be concluded that device-based interaction technique is highly suitable to perform 6 DOF for 3D object manipulation since this technique does not involve the problems stated previously and is proved to improve efficiency and provide high precision in performing the 3D object translation in handheld AR interface [2], [4], [10]. Apparently, this technique provides an intuitive way apart from real hand and fingers interaction technique since users hold the handheld device as a holding tool to handle the 3D object that can enhance the users' AR experience [2], [9], [11]. However, device-based interaction technique may cause registration error when users rotate the virtual object out of the range of view when the angle of the device movement is not relative to the marker placed that should be at the same point of view as the camera and also seen by the users' eyes when the virtual object has been rendered [10]. This problem might cause the roll and pitch axis limitation and also the 360 degree of z-axis (yaw) rotation that require the user to move around, slowing down the virtual object's rotation time [2], [3].

III. 3D OBJECT ROTATION TECHNIQUE IN DEVICE-BASED INTERACTION

Based on the drawbacks discussed in the previous section, we proposed a novel device-based interaction technique that uses device's tilting and skewing amplitudes for 3D object rotation separated from the device movement that has been used for 3D object translation to improve the current device-based interaction technique. This technique uses device tilting and skewing amplitudes to determine the 3D object rotation axes and its direction (clockwise or counter clockwise) while the 3D object is being rotated with an amount of rotation degrees per second automatically after determines its rotation axes and directions to perform complete 3DOF 3D object rotation. Pre-defined device's tilting and sewing amplitudes and also the amount of 3D object rotation degrees per second will be determined in the user study to find out the most suitable device tilting and skewing amplitudes and the 3D object rotation degrees per second for majority users.

A. User Study

In our proposed technique, it involves the determination of the pre-defined device rotation amplitudes before the fully implementation. Thereby, the device tilting and skewing amplitudes will be determined at certain ranges. When the user tilts or skews the device out of the amplitude ranges determined through the pre-defined values, the situation in the first stage been fulfilled, the user will enter the second stage where the 3D object will be rotated with pre-defined amount of 3D object rotation degrees per second.

B. Situational Equation Used in the Proposed Technique

The proposed device-based interaction technique for 3D object rotation is explained in the situational equations as below:

$$\text{IF } (\overleftarrow{P1Q1} > A1)$$
$$\text{IF } (\overrightarrow{P1Q1} > B1) \quad (1)$$
$$\text{IF } (\overleftarrow{P2Q2} > A2)$$
$$\text{IF } (\overrightarrow{P2Q2} > B2) \quad (2)$$
$$\text{IF } (\overleftarrow{P3Q3} > A3)$$
$$\text{IF } (\overrightarrow{P3Q3} > B3) \quad (3)$$

where → and ← represent clockwise and counter clockwise directions;
$P_{\{1,2,3\}}$ represents the device's initial pose (1=roll axis, 2=pitch axis, 3=yaw axis);
$Q_{\{1,2\}}$ represents the new device's pose after tilting ((1=roll axis, 2=pitch axis);
Q_3 represents the new device's pose at yaw axis after skewing;
$A_{\{1,2\}}$ and $B_{\{1,2\}}$ represent the values of tilting amplitudes captured from accelerometer input (1=roll axis, 2=pitch axis)
A_3 and B_3 represent the values of skewing amplitude captured from camera pose input at z axis (yaw)

$$R_{t\{1,2,3\}} = R_{i\{1,2,3\}} + (R_{DEG} \text{ x } T) \quad (4)$$

where R_t represents the new rotation pose of the targeted 3D object after rotation;
R_i represents the new rotation pose of the targeted 3D object before rotation;
R_{DEG} represents the amount of rotation degrees;
T represents real time per second;
$_{t\{1,2,3\}}$ represents the new rotation pose of the targeted 3D object after rotation (1=roll axis, 2=pitch axis, 3=yaw axis);
$_{i\{1,2,3\}}$ represents the initial rotation pose of the targeted 3D object after rotation (1=roll axis, 2=pitch axis, 3=yaw axis);

$$R_t = R_{t\{1,2,3\}} \qquad (5)$$

The final rotation pose of the 3D object represented in R_t is done after the rotation in all axes have been determined ($R_{t\{1,2,3\}}$)

In the equations explained above, the values of A, B and R_{DEG} are pre-defined and will be determined in the user study that may change based on different purposes and applications.

C. Expected Outcome

The expected outcome of the user study that will be carried out are the amplitude values of the device's tilting and skewing poses and also the 3D object's rotation degrees per second. A post-test questionnaire will also be given after the testing section to collect the user's feedbacks about the proposed technique. The feedbacks about the design of the test environment and prototype will also be considered to repair the rotation task representations for further experiments in the future.

D. Experiment Design

This study was designed to obtain the values of the suitable tilting and skewing amplitudes and also the amount of 3D object rotation degrees per second for 3D object rotation. Thus, the task designed in this study is to rotate a scaled 3D object that will be selected randomly by the system from two types of scaled object (Fig. 1) installed; one of the objects represents the engineering object and another represents the natural object. The 3D chair as the engineering object represents geometry objects such as machine building, vehicle and so on including objects that contains repetitive components (like windows, pillars and fixtures) whereas the 3D human as the natural object represents geometry objects such as animals, insects, plants and so on with less repetitive components [13].

(a) (b)

Fig. 1. (a) A 3D virtual chair that represents the engineering object; (b) a 3D human that represent the natural object

The selected 3D object was then displayed on the rendering system together with a wireframe object (Fig. 2a) of the same 3D object displayed after the subject touched the Start button (Fig. 2b) on the device's screen targeted on a printed feature-based AR marker (Fig. 2c). The wireframe object represents the finale rotation pose of the displayed 3D object to be rotated. The position of the 3D object and the wireframe object had been fixed at the middle of the AR marker registered on the device's display screen to avoid the 3D object being moved. A virtual Hold button (Fig. 2d) was also displayed on the bottom left position on the device's screen of view after the subject touched the Start button. The Hold button enables the subject to start the rotation task by touching it, where a green light (Fig. 2e) appears on the right top position on the device's screen of view to show that the 3D object displayed is being hold by the subject. The subject can release the Hold button to unhitch the 3D object.

In the aim of enabling the subject to achieve the finale rotation pose without any miscellaneous since the purpose of this study is to gain feedbacks of the subjects when using the proposed technique not to test the spatial intelligence of the participants, we provided two references for the subjects to determine the finale rotation pose of the displayed 3D object (Fig. 2f).

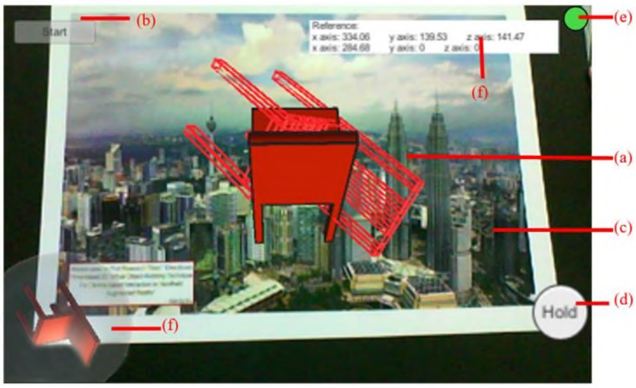
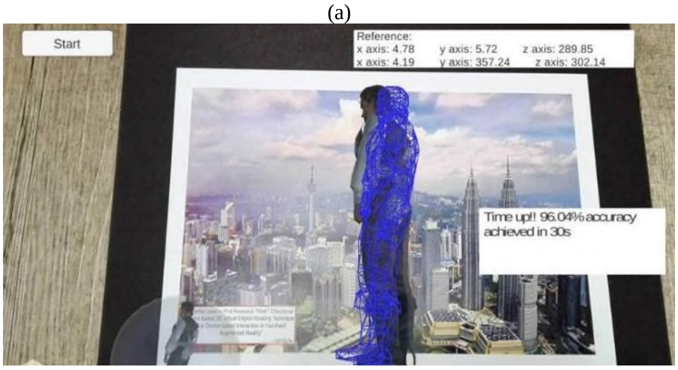


Fig. 2. (a) Wireframe object; (b) Start button; (c) AR marker; (d) Hold button; (e) Green light; (f) References: a visual reference that represents the 3d object in finale rotation pose and a text reference that represents the rotation degrees in 3dof

A countdown of 30 seconds started when the subject touched the Hold button for the first time. The subject can rotate the 3D object displayed when touching the Hold button through device's tilting and skewing actions. When achieving 97 percent accuracy of the rotation degrees within 30 seconds, the Hold button will be terminated (Fig. 3a). Oppositely, if the subject cannot achieve 97 percent accuracy rate, the Hold button will also be terminated after 30seconds (Fig. 3b). The subject can start again the new rotation task by touching the Start button that will be enabled after the Hold button's termination and the new rotation pose will be displayed randomly every time touching the Start button.

(a)



(b)

Fig. 3. (a) When the rotation accuracy achieved 97%, the hold button is terminated; (b) When the time achieves 30 seconds, the hold button is terminated; start button is available when the task stops

E. Apparatus and Workspace

The test was performed in a mobile application built for smartphone and tablet platforms. The application was installed in a low-cost smartphone branded Huawei Y6II Cam-L21 that had been used in this user study for standalone processing and the built-in back-side camera was used to capture live AR video feed. This device had also been determined such that the accelerometer function can perform well for testing purpose. C# and game engine had been used to write the proposed device-based interaction technique for 3D object rotation and the experimental design while C++, OpenGL library had been used to communicate the tracking data to the mobile platform. The device's built-in camera was calibrated and the live view was captured and rendered using OpenCV.

The subjects had been provided with a feature-based AR marker sized A4. The sizes of the scale objects were designed within the range of the marker's frame.

In order to make sure the study's environment is the same for all subjects; the marker had been placed on the floor and the subjects would sit on the floor under the same lighting system facing the AR marker in a comfortable situation with the same or slightly different distances between the subject's eye view and the 3D object displayed. Subjects were allowed to kneel or squat freely if they need to do so.

F. Subjects and Procedure

The study was performed with 10 subjects (7 women and 3 men) recruited from a university with age ranging between 19 and 24 (Mdn = 21.5). The number of subjects is determined after comparing two related studies in similar tasks [11], [12]. We chose the subjects within this age range through the statistic provided by a local smartphone usage report (Nielsen smartphone user segmentation study in year 2015) that shows that the largest group of smartphone users are from this age range. All subjects were reported to be experienced with handheld devices, a total of 5 subjects have AR experiences (3 subjects have intermediate AR experiences and 2 subjects claimed that they are beginners in AR field) and another 5 subjects reported that they are novice users.

At first, the subjects were given a pre-test questionnaire to input their general information, English speaking level and experience level with AR. The proposed technique and the task provided were thoroughly explained to the subjects by a tutor. The subjects were trained and tried to use the proposed technique to rotate the 3D object within 15 minutes in the given test environment. The virtual object rotating task can be carried on several times within 15 minutes.

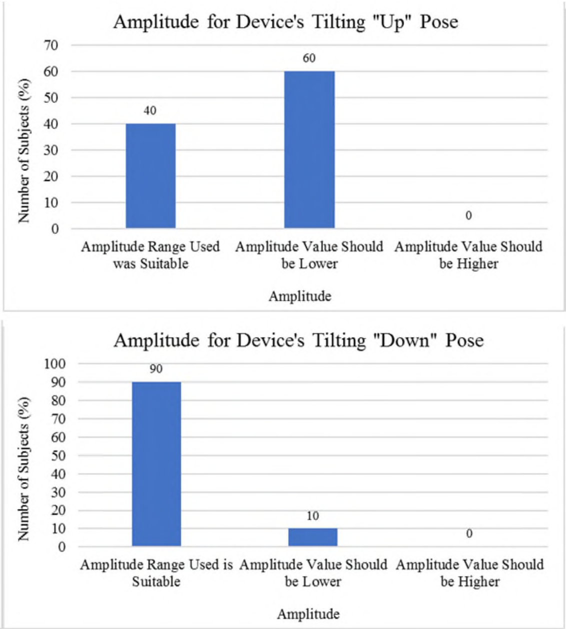
After the testing stage, the subjects were required to give feedbacks according to the proposed technique and the task designed through a post-test questionnaire.

This study aimed to retrieve and determine the pre-define values for the device's tilting and skewing amplitudes for the proposed technique, thus none of the performance variables such as completion time and accuracy rate would be recorded. Further experiments and analysis will be carried out after this study which focus on the performance after the amendment of this prototype based on feedbacks gained from this study is done.

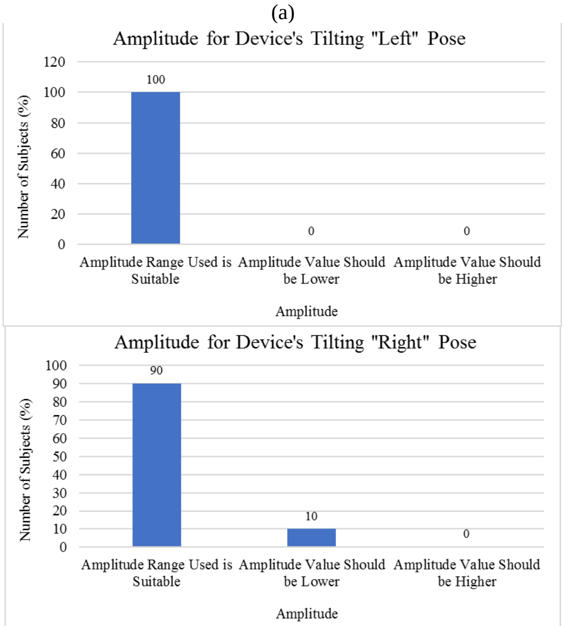
IV. RESULTS AND DISCUSSION

Mode values had been used to determine the device's tilting and skewing amplitude and also the amount of 3D object rotation degrees per second. We used the Fisher-Freeman-Halton test [14] (for small sample size) to evaluate whether AR experiences affect the subjects' choices for the device's tilting and skewing amplitude and also the amount of 3D object rotation degrees per second by testing both independence and homogeneity. Besides, some core aspects of the proposed technique implementation: device pose, the position of built-in camera, single-hand or both-hands holding action would be verified to support previous studies that proved the device-based interaction technique can be implemented by using both hands and so the proposed technique; the device is usually used in landscape orientation for AR interface or portrait and the position of built-in camera is usually rotated to be located on the left-hand side or right-hand side for the users. Some design' elements of this experiment had also been determined either to be eliminated or retained through the mode value for each category.

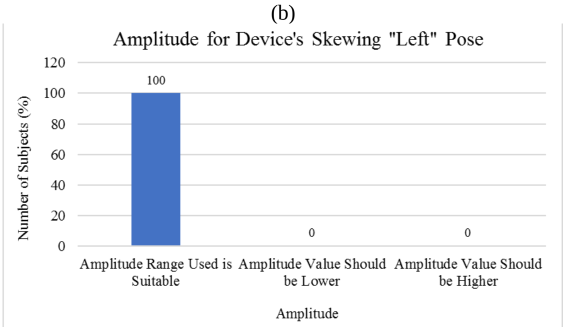
We evaluated all the data. Fig. 4 illustrates the mode which represents the most chosen device's tilting and skewing amplitudes together with the most chosen value for the amount of 3D object rotation degrees per second.

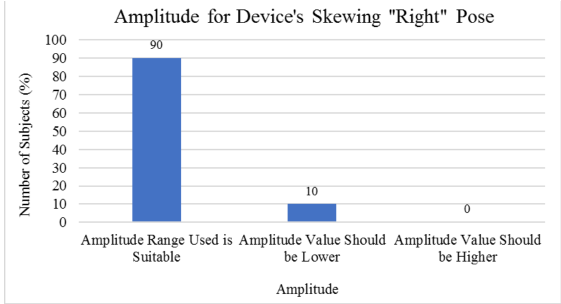


(a)

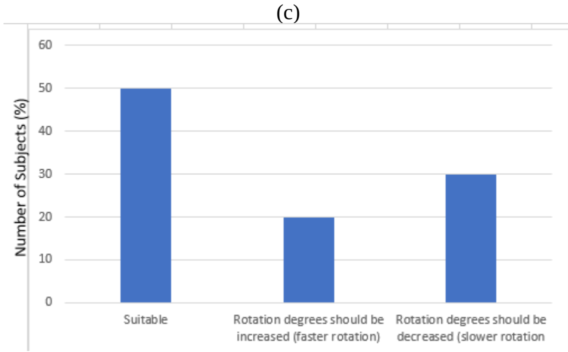


(b)





(c)



(d)

Fig. 4. (a) The most subjects' chosen amplitudes for device's tilting "up" and "down" poses; (b) The most subjects' chosen amplitudes for device's tilting "left" and "right" poses; (c) The most subjects' chosen amplitudes for device's skewing "left" and "right" poses; (d) The most subject's chosen amount of 3d object rotation degrees is suitable, is fast or is slow

As seen from the results shown in Fig. 4, most of the subjects agree with the amplitude values pre-defined in the proposed technique for device's tilting and skewing poses. However, for device's tilting "up" pose, most of the subjects prefer to lower the pose for 3D object's rotation around pitch axis at counter clockwise direction. For the amount of 3D object's rotation degrees, some subjects prefer to increase the rotation degrees for a faster rotation and some of them prefer decrease the rotation degrees for a slower rotation. Though so, still half of the subjects (50%) prefer the pre-defined amount of 3D object's rotation degrees per second used in this prototype.

Thus, we decided to retain the pre-defined amplitudes range and 3D object's rotation degrees per second used for the proposed virtual object rotating technique by only changing the amplitude value for the device's tilting "up" pose based on the subjects feedbacks. The situational equation below showed the pre-defined values determined through this study:

$$\text{IF } (\overleftarrow{P_1Q_1} > -0.12) \quad \text{IF } (\overrightarrow{P_1Q_1} > 0.12) \tag{6}$$

$$\text{IF } (\overleftarrow{P_2Q_2} > -0.05) \text{ (this value would be changed to } \overleftarrow{P_2Q_2} > -0.1 \text{ based on the study's feedbacks)} \quad \text{IF } (\overrightarrow{P_2Q_2} > -0.3) \tag{7}$$

$$\text{IF } (\overleftarrow{P_3Q_3} > -0.1) \quad \text{IF } (\overrightarrow{P_3Q_3} > 0.1) \tag{8}$$

Equation 6, 7 and 8 is relevant to equations 1, 2 and 3 stated in the previous section.

$$R_{t\{1,2,3\}} = R_{i\{1,2,3\}} + (23.5° \times T) \quad (9)$$

Equation 9 is relevant to equation 4 stated in the previous section.

For the independence and homogeneity tests, we found out that there is no dependence between the AR experience and the subject's choices towards device's tilting and skewing pose and also the 3D object rotation degrees per second (Table 1). The distribution of the subjects with or without AR experiences are exactly the same for each choice for device's tilting and skewing amplitudes and also the amount of 3D object rotation degrees per second. Thus, we believe that AR experiences do not affect the subjects' choices towards the device's tilting and skewing amplitudes and the amount of 3D object rotation degrees per second.

TABLE 1. FISHER-FREEMAN-HALTON TEST (FOR SUBJECTS WITH AR EXPERIENCES AND WITHOUT AR EXPERIENCES)

Questionnaire's Items	**Fisher Exact Probability (Two-tailed)**	**Dependence**
Device's tilting "up" pose	P=0.286 (ns)	No
Device's tilting "down" pose	P≈1.000 (ns)	No
Device's tilting "left" pose	P=1.000 (ns)	No
Device's tilting "right" pose	P≈1.000 (ns)	No
Device's skewing "left" pose	P=1.000 (ns)	No
Device's skewing "right" pose	P≈1.000 (ns)	No
Virtual object's rotation (degrees per second)	P≈1.000 (ns)	No

Note: ns = not significant; * significant at alpha =0.05; ** significant at alpha =0.01

V. CONCLUSION

This paper has presented a user study designed to perform 3D object rotation by using the proposed novel device-based interaction technique explained in section 4 base on the concept of the existing device-based interaction, which integrates the device's tilting and skewing amplitudes for 3D object rotation in device-based interaction in handheld AR interface.

Based on the feedbacks gained from the user study, the pre-defined value stated in Eq. 7 (P_2Q_2) had been changed to >-0.1 while the others pre-defined values remain unchanged. Some aspects of the experiment designed had also been changed based on the participants' feedback. We had fixed the application to landscape left mode since all subjects hold the smartphone in the landscape left orientation.

We proposed a separated input types for 3D object translation and rotation by considering the drawbacks caused by the current device-based interaction technique that uses device movement for both actions (translation and rotation) although it had been proven that the device movement is not suitable to be used for 3D object rotation.

Based on the feedbacks given by some of the subjects, they are more familiar with 3D chair rather than a 3D human which are conformed with the anecdotal reports that suggest the use of a everyday, more familiar, and less symmetrical object, such as a chair, could reduce the perceptual complexity of the docking task while rotation task is a part of it and also the goal of the rotation task in this experiment is to evaluate the interaction technique and not the spatial intelligence of the participants, we decided to conducted the upcoming evaluation comparing both techniques (the proposed technique and the existing technique) via performances (task completion time and accuracy) by only using the 3D chair following the previous studies [12], [15]. Besides, only using one type of 3D object might avoid the latency effect caused by different 3D objects with different polygonal complexities.

The technique presented in this paper uses different inputs for 3D object rotation thus avoiding problems due by using only device movement for both translation and rotation. An accidentally movement of the 3D object caused by the device movements may be avoided through the usage of our proposed technique. Thus, a separated user study also had been carried out to evaluate the performances of our proposed technique in performing 3D object manipulation tasks (both translation and rotation) integrally by differentiate the 3D object translation and rotation through interaction techniques.

ACKNOWLEDGMENT

Thanks go to other members of the Virtual, Vision and Visualization Research Lab (UTM ViCubeLab), Universiti Teknologi Malaysia for their time, effort, and enthusiasm.

REFERENCES

[1] D. Bowman, E. Kruijff, J. LaVoila, and I. Poupyrev. (2005). *3D User Interfaces: Theory and Practice.* United States: Adisson-Wesley.

[2] A. Mossel, B. Venditti, and H. Kaufmann. (2013). 3D Touch and HOMER-S: Intuitive Manipulation Techniques for One-Handed Handheld Augmented Reality. *Proceedings of the Virtual Reality International Conference: Laval Virtual, ACM.*

[3] A. Marzo, B. Bossavit, and M. Hachet. (2014). Combining Multi-Touch Input and Device Movement For 3D Manipulations In Mobile Augmented Reality Environments. *Proceedings of the 2nd ACM Symposium on Spatial User Interaction*, 13-16.

[4] C. F. Yusof, H. Bai, M. Billinghurst, and M.S. Sunar. (2016). A Review of 3D Gesture Interaction for Handheld Augmented Reality. *Jurnal Teknologi.* 78(2-2).

[5] E. S. Goh, M. S. Sunar, and A. W. Ismail. (2020). Device-Based Manipulation Technique with Separated Control Structures for 3D Object Translation and Rotation in Handheld Mobile AR. *International Journal of Human-Computer Studies*, 141. doi.org/10.1016/j.ijhcs.2020.102433.

[6] E. S. Goh, M. S. Sunar, and A. W. Ismail. (2018). An Inertial Device-based User Interaction with Occlusion-free Object Handling in a Handheld Augmented Reality. *International Journal of Integrated Engineering*. 10(6), 159-168.

[7] W. Hürst, and C.V. Wezel. (2013). Gesture-Based Interaction Via Finger Tracking for Mobile Augmented Reality. *Multimedia Tools and Applications*, 62(1), 233-258.

[8] M. Hancock, S. Carpendale, and A. Cockburn. (2007). Shallow-Depth 3D Interaction: Design and Evaluation of One-, Two- and Three-Touch Techniques. *Proceedings of the 2007 ACM*

Conference on Human Factors in Computing Systems, 1147-1156.

[9] A. Martinet, G. Casiez, and L. Grisoni. (2012). Integrality and Separability of Multitouch Interaction Techniques in 3D Manipulation Tasks. *IEEE Transactions on Visualization and Computer Graphics,* 18(3), 369-380.

[10] A. Henrysson, J. Marshall, and M. Billinghurst. (2007). Experiments in 3d Interaction for Mobile Phone AR. *Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia, SER. GRAPHITE '07, ACM,* 187-194.

[11] T. Tanikawa, H. Uzuka, and T. Narumi, and M. Hirose. (2015). Integrated View-Input AR Interaction for Virtual Object Manipulation Using Tablets and Smart Phones. *Proceedings of the 2015 International Conference on Advances in Computer Entertainment Technology, ACM.*

[12] A. Samini, and K. L. Palmerius. (2016). A Study on Improving Close and Distant Device Movement Pose Manipulation for Hand-Held Augmented Reality. *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology,* 121-128.

[13] A. A. Globa, M. Donn, and O. A. Ulchitskiy. (2016). Metrics for Measuring Complexity of Geometric Models. *Scientific Visualization*, 8(5), 74-82.

[14] G.H. Freeman, and J.H. Halton. (1951). Note On Exact Treatment of Contingency, Goodness of Fit and Other Problems of Significance. *Biometrika*, 38, 141-149.

[15] V. Vuibert, W. Stuerzlinger, and J. R. Cooperstock. (2015). Evaluation of Docking Task Performance Using Mid-air Interaction Techniques. *Proceedings of the Conference on Spatial User Interaction (SUI),* 44-52.