Head Mounted Display Interaction Evaluation: Manipulating Virtual Objects in Augmented Reality

Maite Frutos-Pascual[1], Chris Creed[1], and Ian Williams[1]

DMT Lab, School of Computing and Digital Technology,
Birmingham City University, United Kingdom
`maite.frutos@bcu.ac.uk`

Abstract. Augmented Reality (AR) is getting close to real use cases, which is driving the creation of innovative applications and the unprecedented growth of Head-Mounted Display (HMD) devices in consumer availability. However, at present there is a lack of guidelines, common form factors and standard interaction paradigms between devices, which has resulted in each HMD manufacturer creating their own specifications. This paper presents the first experimental evaluation of two AR HMDs evaluating their interaction paradigms, namely we used the HoloLens v1 (metaphoric interaction) and Meta2 (isomorphic interaction). We report on precision, interactivity and usability metrics in an object manipulation task-based user study. 20 participants took part in this study and significant differences were found between interaction paradigms for translation tasks, where the isomorphic mapped interaction outperformed the metaphoric mapped interaction in both time to completion and accuracy, while the contrary was found for the resize task. From an interaction perspective, the isomorphic mapped interaction (using the Meta2) was perceived as more natural and usable with a significantly higher usability score and a significantly lower task-load index. However, when task accuracy and time to completion is key mixed interaction paradigms need to be considered.

Keywords: Augmented Reality, Hand Interaction, Natural Interaction

1 Introduction

Augmented Reality (AR) is defined as an interactive technology that incorporates virtual objects into the real world [30]. AR is now maturing and getting close to real use cases [35], this leading to the creation of innovative applications [38] and the unprecedented growth of Head-Mounted Display (HMD) devices and consumer availability. Similarly, User Interfaces are rapidly evolving beyond traditional desktop and mobile environments. Technologies such as AR, tangible interfaces and immersive displays currently offer more natural ways of interaction, leveraging user interaction capabilities with the real world [7].

This timely combination of hardware availability and technology advances has led to the incorporation of AR in real application cases outside laboratory

environments, leading it to been dubbed as one of the key technologies in the Industry 4.0 revolution [27, 33]. Its current use-cases have expanded beyond to health-care [24], education [3], and tourism [54], among others. This growth of AR applications outside traditional Human Computer Interaction (HCI) research environments implies that technology is currently being used outside laboratory environments. As such, there is an urgent need to understand how people naturally interact with immersive systems to make sure AR can reach its full potential. However, these emerging technologies currently have no established design guidelines, interaction metaphors or form factors [21], which has resulted in each HMD manufacturer creating their own.

This lack of guidelines and standard interaction paradigms is evident when analysing two of the most current AR HMDs enabling freehand interaction, that is without wearables or controllers, applications and research, namely the Microsoft HoloLens v1[1] and Meta's Meta2[2] HMDs. Both devices currently offer direct freehand interaction [20], however each supports different interaction paradigms and form factors that may have an effect on their perceived usability (i.e. the position of AR objects, the headset Field Of View (FOV) and also the recommended user interaction space). In terms of the inherent interaction paradigms, the Microsoft HoloLens v1 offer a head/gaze and tap based interaction paradigm (see Fig. 1(b)) which is a restricted form of gesture interaction [29] that could be labelled as a metaphoric mapping interaction [37], as it is similar to single mouse clicks on a desktop. The Meta2 offers a spatially co-located interaction [29] where users can manually (see Fig. 1(d)) or bi-manually (see Fig. 1(e)) interact with virtual objects and it is somewhat comparable to manipulating objects in real life, thus it was labelled as isomorphic mapping interaction, defined as a one-to-one literal spatial relation between actions and system effects [37]. These interaction paradigms offer a distinctive natural interaction, one based on remote manipulation of objects akin to current PC interaction paradigms (metaphoric) while the other mimics real world tasks, enabling a physical manipulation (isomorphic).

While both interaction paradigms and device form factors offer viable interaction methods [29], usability, interactivity and precision of these two devices for task-based AR scenarios has not been fully addressed before. Therefore no formal studies have aimed to evaluate a user's response to these devices and their inherent interaction paradigms in AR. Research in this area can help identifying users' real needs and expectations before designing application concepts [8]. We present the first study evaluating two commercially available AR devices and their inherent interaction paradigms for AR object manipulation. These paradigms are applied following the stringent design considerations of the competing devices. We evaluate these on the users' ability to manipulate virtual objects. We measured effectiveness, efficiency and user satisfaction for each condition and task.

---

[1] https://www.microsoft.com/en-us/hololens, *(12th of December 2018)*
[2] https://www.metavision.com/ *(12th of December 2018)*

Finally we present our findings and considerations, thus enabling future AR developers and UX designers to determine which is the most suitable device and interaction paradigm to fit their interaction and user experience needs.
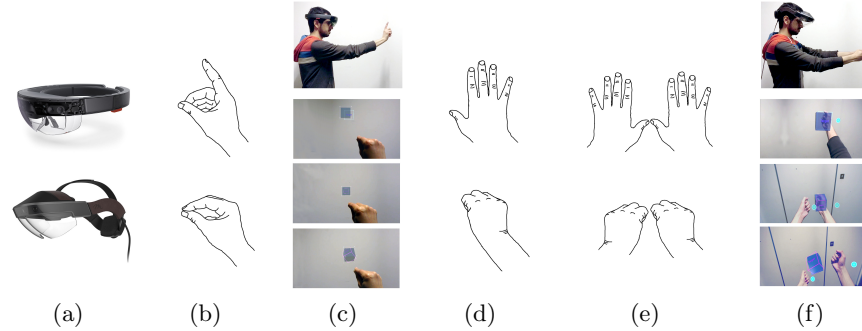

(a)   (b)   (c)   (d)   (e)   (f)

Fig. 1. Microsoft HoloLens v1 1(a) supports metaphoric interaction via the head/gaze and tap 1(b) 1(c). Meta2 1(a) supports a isomorphic interaction in both Manual 1(d) 1(f) and BiManual configurations 1(e) 1(f).

2   Related Work

The accessibility of consumer AR systems is growing and the adoption of AR systems into application domains is becoming commonplace [14]. With the recent advances in consumer hardware and the integration of on-board sensor technology, interactive AR applications in healthcare [15, 18], education [13] and industry [19] are now possible. This advanced hardware is driving the integration of AR into the workplace with many industries keen on applying AR into their work-flow processes [22]. Now naive users have the possibility to use AR systems that contain plausible realistic environments and also facilitate direct real-time interaction between themselves and the virtual objects in ways that were not easily possible previously.

2.1   Commercial Devices

Commercial head-mounted see-through AR solutions offer new platforms for AR developers, designers and industry applications through facilitating an increased ease of development. However they all differ in their fundamental hardware configurations (e.g tethering, tracking and physical dimensions) and conversely offer different paradigms interaction with AR content (e.g. metamorphic as with the HoloLens, isomorphic as with the Meta2 and the use of hand-held controllers as in the Magic Leap One). As new devices become available in the near future, it is crucial to understand user's response to available devices and form factors, so new cross-device interaction standards can be established.

2.2   Interaction Approaches

Research studies focused around AR have explored and compared a range of different interaction paradigms to examine their suitability to support AR experiences. Lee *et al.* [34] compared the usability of multi-modal interfaces (i.e. gestures and speech recognition) against speech-only, and gesture-only approaches (across a series of object manipulation tasks). Similarly, the use of gestures was explored for direct physical grasping in comparison to the combination of speech and gestures for indirect interactions [41]. While much research has studied the application of freehand interaction in AR, most have focused on comparing freehand gestures alone [45]. This work has shown that overall a gesture based interaction paradigm can offer ease of access and naturalness to users, and does deliver an intuitive and effective interaction [26, 36, 40, 43, 52].

Researchers are currently evolving gesture interactions into physical interactions, thus, completely mimicking the interaction performed in the real world, into AR environments [46, 49]. The work of Al-Kalbani *et al* [4] and [5] have sought to understand the complex nature of the human grasp for freehand interaction and defined the problems associated with grasping objects in AR, Ciodota *et al* [15] evaluated the application of grasps in AR and also Swan *et al* [51] evaluated the problems associated with representing and reaching the true located space of augmented objects.

While different input combination approaches have been explored in the literature, together with different gesture based interactions and the impact of physical interactions alone. There has since been limited research focused on evaluating the problems and benefits to compare different interaction paradigms inherent to off-the-shelf commercial devices.

3   Methods

An experiment was conducted that analysed two commercially available AR devices with their underlying interaction paradigms. We used a within participants design. As this study focuses on the comparison of two off-the-shelf AR HMDs, we used the interaction paradigms recommended by each device manufacturer and their stringent environment configuration as deemed to be the most suitable and stable way of interfacing with the hardware. As this study focuses on the evaluation of devices form factors, design guidelines and interaction metaphors on perceived usability of naive participants, the use of configurations outside the recommended guidelines in [1, 2] were not included in this study.

3.1   Conditions

Two different interaction paradigms were studied in this work. Interaction paradigms outlined in this section are the recommended for optimal object manipulation by the device manufacturers [1, 2]. Both conditions facilitate direct freehand interaction between the user and the AR content. However, they offer a different

interaction approach that may become apparent when interacting virtual objects in an AR environment. Tasks described in subsection 3.4 were performed in each of these conditions. There is no current knowledge of which interaction paradigm is preferred for object manipulation in AR environments.

- **Metaphoric mapped interaction** *(using the HoloLens v1).* Following Macaranas *et al.* definition, this is defined as the mental models generated from repeated patterns in everyday experiences [37]. Microsoft HoloLens v1 hand interaction paradigm replicates the most classic input paradigm in computer systems, the point and click interaction using the mouse. Head gaze control is used for targeting virtual objects (commonly referred to as *holograms*) in the 3D space and tap gesture interaction is for acting on specific selected virtual objects (i.e. similar to mouse clicks) [2]. This offers a remote interaction paradigm, enabling users to access any virtual object in the interaction space remotely.
- **Isomorphic mapped interaction** *(using the Meta2).* This interaction is defined as the one-to-one literal spatial relations between input actions and their resulting system effects [37]. Meta2 offers this level of interaction, requiring the user to be physically placed in the location of the virtual object to be able to access it and trigger the interaction. This creates a spatial equivalence between the virtual environment and user's actions which is akin to manipulating virtual objects as if they were real objects [29, 42].

3.2   Apparatus

We built a custom experimental framework for the HoloLens v1 and Meta2 using Unity 2017.3, Windows Holographic platform[3] and the Mixed Reality Toolkit[4] for the HoloLens v1, and Meta2 SDK 2.7 [5] for the Meta2. C# was used as scripting language.

- **Microsoft HoloLens v1.** The HoloLens v1 is a wireless, optical see-through stereographic AR HMD with a 30×17 Field of View (FOV). The device has built-in integrated Inertial Measurement Uni sensor, a depth camera, four environment understanding cameras, HD video camera enabling Mixed Reality capture and four microphones. It supports voice input, gesture recognition and head/gaze positioning tracking. It weights 579g, as per its specification description.
- **Meta's Meta2.** The Meta2 is a tethered, optical see-through stereographic AR HMD with a 90 degrees FOV. The device has built-in hand interaction and positional tracking sensors, a 720p front-facing RGB camera, four surround speakers and three microphones. It natively supports hand tracking and head/gaze positioning tracking. It weights 500g, as per its specification description.

---

[3] https://developer.microsoft.com/windows/mixed-reality, *(12th of December 2018)*
[4] https://github.com/Microsoft/MixedRealityToolkit-Unity, *(12th of December 2018)*
[5] https://devcenter.metavision.com/ , *(12th of December 2018)*

(a)   (b)   (c)   (d)
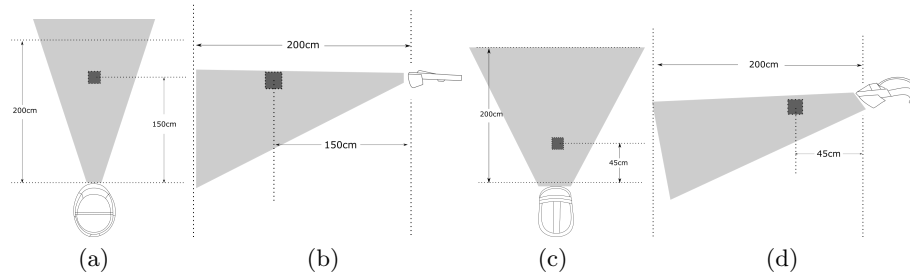
Fig. 2. Microsoft HoloLens v1 - Interaction space configuration. 2(a) Top view with the object displayed 150 cm away from the user standing position. 2(b) Side view with the interactive virtual object displayed 150 cm away and in the horizon line of the device. **Meta Meta2** - Interaction space configuration. 2(c) Top view with the object displayed 45 cm away from the user standing position.2(d) Side view with the interactive virtual object displayed 45 cm away and 10° below the horizon line of the device.

3.3   Origin position

Interactive virtual objects were always presented at the initial origin position. This position was selected to comply with device manufacturers' recommendations as displayed in Figure 2.

– **Metaphoric mapped interaction** *(using the HoloLens v1)***:** Interactive virtual objects and targets for this condition were placed within the preferred interacting area as reported in the *HoloLens v1 Development Guidelines - Holograms* [2]. The initial position of the interactive virtual object was at the device's horizon line and 150 cm away from the device, as reported in Figures 2(a) and 2(b) and in alignment with manufacturer's hologram placement guidelines [2].
– **Isomorphic mapped interaction** *(using the Meta2)***:** The interactive virtual object and targets were placed within the preferred interacting area as reported in the *Meta2 User Interface Guidelines and Ergonomics* [1]. The initial position of the interactive virtual object was 10 degrees below the device's horizon line and 45 cm away from it, within arm reach, as reported in Figures 2(c) and 2(d) and in alignment with their guidelines about the optimal interaction space [1].

Different origin positions were used to comply with manufacturers' recommendations and devices form factors; being the Meta2 a tethered device that enable co-located interaction, manufacturer's guidelines recommend virtual objects to be placed within arm reach [1] while the Microsoft HoloLens v1 guidelines recommend a remote interaction, with the virtual objects to be placed between 1.25m and 5m away, with the ideal interaction space being 1+ m away [2].

3.4   Tasks

Tasks were selected based on the survey presented in [42], where Piumsomboom *et al.* categorised the most common AR interaction tasks into 6 categories. This study analysed a subset of tasks extracted from the Transform category complemented by an additional move in depth (z axis) task.

For each task the user is presented with an interactive object (a cube) and a target object. The target object represents the different task attributes (i.e. position for move tasks, size for re-size task and rotation orientation for the rotate task). Target objects were not interactive and they were displayed as virtual outlined objects using different colour and texture to avoid confusion. Each task was comprised of four different targets appearing one after the other in random order. The interactive object resets to its original position, rotation and size after every target.

– **Move Tasks.** Participants were asked to move a virtual interactive cube in the 3D space to match the position of a target. Three different move tasks were performed:
- **Short distance.** In this tasks targets appeared within the Field of View (FOV) of the devices. Table 1 shows the specific position of the targets while Figure 3(a) displays a graphical representation of their spatial distribution.
- **Long distance.** Targets appeared spatially distant to the interactive object initial position for both conditions in 3.1. These targets will appear outside the initial FOV of the devices. Table 1 shows the specific position of the targets while Figure 3(a) displays a graphical representation of their spatial distribution.
- **Depth (z axis).** Depth perception and estimation is a long-standing research field in Augmented Reality literature, with a well known set of perception issues and estimation problems defined and studied [16, 50, 28, 48, 44, 32, 31]. We considered relevant to extend Piumsomboom *et al.* [42] move task category with an additional movement solely in z (depth). This tasks is formed by four targets that appeared one after the other spatially distributed in the interaction space, with two of them rendering virtually close to the user initial standing position and the other two further away, as described in Table 1 and Figure 3(b).

– **Resize Task.** Participants were asked to uniformly resize a virtual interactive cube in the 3D space to match the size of a target. The task was to grow or shrink a 10 cm width x 10 cm height x 10 cm depth virtual cube to different target sizes. Target size distribution is shown in Figure 4 and target sizes are reported in Table 1.
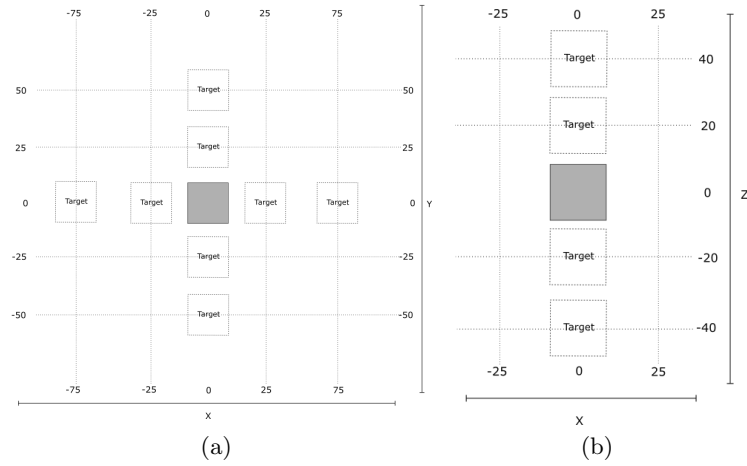
Fig. 3. Spatial distribution of targets for move tasks. 3(a) shows target distribution for tasks move short (inner targets) and move long (outer targets). 3(b) displays targets for move depth task. Targets are shown as distance from the origin in cm. The interactive virtual object initial position is represented as a dark gray square in 3(a) and 3(b).
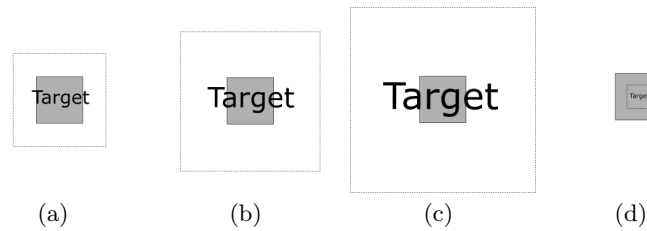

Fig. 4. Resize targets for the interactive object (gray square). 5(a) Target1 in Table 1; 5(b) Target2 in Table 1; 5(c) is Target3 in 1 and 5(d) is Target4 in Table 1.

- **Rotate Task.** Participants were asked to rotate a virtual cube in the 3D space to match the rotation of a target. Target distribution is displayed in Figure 5 and specific target rotation values are reported in Table 1; covering roll (x axis rotation), pitch (y axis rotation) and yaw (z axis rotation) as described by [42] and a combination of the three (x, y and z axis rotation).

3.5   Environment

Participants performed the test in a controlled environment under laboratory conditions. A clear room layout was used. Interactive space dimensions were 270 cm by 180 cm. The test room was lit by a 2700k (*warm white)* fluorescent with controlled external light source. Task initial position was marked on the floor and participants were free to move around the space to complete the target tasks.
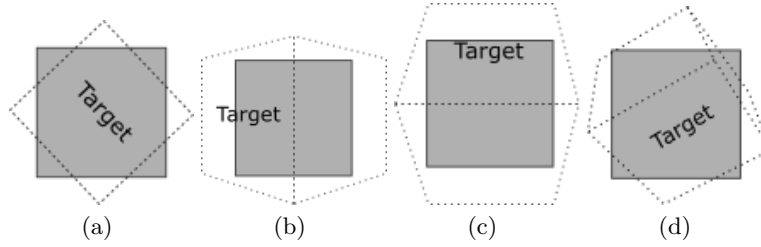
Fig. 5. Rotate targets for the interactive object (gray square). 5(a) Target1 in Table 1; 5(b) Target2 in Table 1; 5(c) is Target3 in 1 and 5(d) is Target4 in Table 1.

Table 1. Target distribution per tasks. Targets are displayed as distance away from the initial position for move tasks. Positions refer to distance from the origin [in cm] in *x, y z*. Targets are displayed as target size [in cm] for resize task. The same target resize options were displayed in both conditions. Original size of the interactive object was *(10 cm × 10 cm × 10 cm)*. The same target rotation options were displayed in both interaction conditions. Rotation displayed per axis (x, y, z) in degrees. Initial rotation of the interactive object was $(0°, 0°, 0°)$.

Move - Short distance [in cm]			
Target1	Target2	Target3	Target4
(25, 0, 0)	*(-25, 0, 0)*	*(0, 25, 0)*	*(0, -25, 0)*
Move - Long distance [in cm]			
Target1	Target2	Target3	Target4
(75, 0, 0)	*(-75, 0, 0)*	*(0, 50, 0)*	*(0, -50, 0)*
Move - Depth [in cm]			
Target1	Target2	Target3	Target4
(0, 0, -40)	*(0, 0, -20)*	*(0, 0, 20)*	*(0, 0, 40)*
Resize [in cm]			
Target1	Target2	Target3	Target4
(2×)	*(3x)*	*(4×)*	*(0.5×)*
20×20×20	30×30×30	40×40×40	5×5×5
Rotate [in degrees]			
Target1	Target2	Target3	Target4
$(45°, 0°, 0°)$	$(0°, 45°, 0°)$	$(0°, 0°, 45°)$	$(45°, 45°, 45°)$

3.6   Participants

20 right-handed participants (5 female, 15 male) from a population of university students and staff members were recruited to take part in this study. Participants' mean age was 28.05 (SD: 9.78).

All participants performed the tasks described in section 3.4 in both conditions. Participants completed a standardised consent form and were not compensated. Visual acuity of participants was measured using a Snellen chart, each

participant was also required to pass an Ishihara test to exclude for colour blindness. Participants with colour blindness and/or visual acuity of < 0.80 (where 20/20 is 1.0) were not included in this study.

Participants were asked to self-assess their level of experience with Mixed and Augmented reality systems, with 12 participants reporting being *novices* with the technology and 8 having an *average* knowledge. 6 participants reported on having previous experience with HMDs, notably they reported on having used Virtual Reality headsets (i.e. Oculus Rift, Samsung Gear or Google Cardboard). None of the participants had any previous substantial experience with HMDs in Augmented Reality.

3.7   Protocol

A within participants test protocol design was used. Participants were counterbalanced across the two interaction paradigms; half of the participants (10) started the tests with the metaphoric mapped interaction using the Microsoft HoloLens v1 and the other half (10) started with the isomorphic mapped interaction approach using the Meta2.

a) **Pre-test.** Prior to the test, participants were given a written informed consent where the test protocol and main aim of the study was described. Additionally, participants filled in a pre-test questionnaire inquiring about their background level of experience with AR systems and their previous use of HMDs (if any).

b) **Calibration.** Before each test, the test coordinator followed manufacturers' guidelines to help participants fitting the devices in the most suitable and comfortable way. For the Meta2, the additional headset calibration process was launched as per the manufacturer's recommendation.

c) **Training.** Participants underwent initial hand interaction and task training with both devices.
  – **Hand Interaction training:** Participants were trained with the interaction paradigm inherent to the device. For this, the standard built-in gesture/hand interaction training application provided by the devices' manufacturers were used. For the HoloLens v1, the Microsoft *Learn Gesture App* was used to help users understand the interaction paradigm. For the Meta2, the *Hand Training App* provided with the Meta2 2.5 Beta SDK was used to help users understand the isomorphic interaction required for this condition.
  – **Task training:** Once participants were comfortable with the device and the hand tracking, recognition and interaction system, they were asked to undergo specific task-related training. This task training was the same for both conditions in 3.1. Participants were trained in a representative version of tasks in 3.4: move, rotate and resize, one after the other in this order.

d) **Test.** Once participants were comfortable with the interaction paradigm, the hand recognition system and user interface, participants were presented with

the main experimental task. Tasks reported in 3.4 were loaded in random counterbalanced order, as with the different targets in each task. Participants were free to move around the interaction space. In between tasks and targets, they were asked to get back to the starting position, marked clearly on the floor.

e) **Post-test.** After each of the conditions were completed, participants were asked to fill in the NASA TLX [23], the System Usability Scale (SUS) [11] and a post-test questionnaire asking about their experience with the device and their opinion on device specifics (FOV, tracking accuracy, hand tracking accuracy, preferred interaction).

3.8   Metrics

Following Piumsomboon *et al.* [41] definitions, we define the interactivity as the users' ability to interact with the virtual objects and the precision as the level of control the user has when interacting. An example measure of precision would be how accurately the user can rotate or move an object to match a target, while we reported on interactivity as the perceived usability of the system.

- **Precision metrics.** Precision metrics were defined as follows: *time to completion* as the time it took to complete the task and *accuracy* as the difference between the target and the interactive object at the end of the task. For move tasks, the accuracy is measured as the euclidean distance between the target and the interactive object at the end of the task; for resize it was the difference between the target scale and the interactive object scale at the end of the task; for rotation it was the difference between the total rotation of all angles between the target cube and the interactive object. Angles were normalised to a $(45°)$ degree range due to the nature of the interactive object being a cube with the same colour surfaces.
- **Interaction metrics.** Interaction metrics were defined as the subjective metrics obtained from users using observation, NASA TLX [23], the System Usability Scale (SUS) [11] and a post-test questionnaire.

3.9   Hypotheses

Following the methodology defined in this paper, we propose the hypotheses listed below:

- **Hypothesis H_1.** The interaction paradigm within the device conditions (see 3.1) has no effect on time to completion for tasks in 3.4.
  - *Alternative hypothesis AH_1.* The interaction paradigm within the device conditions used has an effect on time to completion for tasks in 3.4.
- **Hypothesis H_2:** The interaction paradigm within the device conditions (see 3.1) has no effect on accuracy for every task reported in 3.4.
  - *Alternative hypothesis AH_2.* The interaction paradigm within the device conditions has an effect on accuracy for every task reported in 3.4.

– **Hypothesis H$_3$:** The interaction paradigm within the device conditions (see 3.1) has no effect on the interaction metrics; especially in *a)* NASA-TLX and *b)* SUS
  - *Alternative hypothesis AH$_3$.* The interaction paradigm within the device conditions has no effect on the interaction metrics reported in metrics; especially in *a)* NASA-TLX and *b)* SUS

3.10   Statistical Analysis

The Shapiro-Wilk [47] normality test found the data to be non-parametric and not normally distributed. We test for significance between the two conditions described in section 3.1 for the metrics shown in 3.8 using a non parametric Wilcoxon signed-rank test [53] with an alpha of 5%.

4   Results

4.1   Precision metrics

– **Move - Short distance.**
  - *Time To Completion.* A statistically significant difference in completion time was found between the two interaction conditions in 3.1 ($Z = 1.10 \times 10^3, p < 0.05$). Hypothesis **H$_1$** was rejected for this task. Therefore we accepted alternative hypothesis **AH$_1$**. Time to completion for this task is displayed in Figure 6(a).
  - *Accuracy:* A statistically significant difference was found when comparing the average euclidean distance between conditions ($Z = 1.15 \times 10^3, p < 0.05$). Hypothesis **H$_2$** was rejected for this task. Therefore, we accepted the alternative hypothesis **AH$_2$**. Overall accuracy and time to completion per task are displayed in Table 2.
– **Move - Long distance.**
  - *Time to Completion.* A statistically significant difference in completion time was found between the two interaction conditions in 3.1 ($Z = 1.07 \times 10^3, p < 0.05$). Hypothesis **H$_1$** was rejected for this task. Therefore we accepted alternative hypothesis **AH$_1$**. Time to completion for this task is displayed in Figure 6(b).
  - *Accuracy.* A statistically significant difference was found when comparing the average euclidean distance between conditions in 3.1 ($Z = 1.13 \times 10^3, p < 0.05$). Hypothesis **H$_2$** was rejected for this task. Therefore we accepted alternative hypothesis **AH$_2$** for this task. Overall accuracy and time to completion per task are reported in Table 2.
– **Move - Depth.**
  - *Time to Completion.* A statistically significant difference in completion time was found between the two conditions in 3.1 ($Z = 1.12 \times 10^3, p < 0.05$). Hypothesis **H$_1$** was rejected for this task. Therefore we accepted alternative hypothesis **AH$_1$**. Time to completion for this task is displayed in Figure 6(c).

- *Accuracy.* A statistically significant difference was found when comparing the average euclidean distance between conditions ($Z = 1.13 \times 10^3, p < 0.05$). Hypothesis $\mathbf{H_2}$ was rejected for this task. Therefore we accepted alternative hypothesis $\mathbf{AH_2}$ for this task. Overall accuracy and time to completion per task are presented in Table 2.
- **Resize.**
  - *Time to Completion.* A statistically significant difference in completion time was found between the two interaction conditions in 3.1 ($Z = 1.14 \times 10^3, p < 0.05$). Therefore, hypothesis $\mathbf{H_1}$ was rejected for this task and alternative hypothesis $\mathbf{AH_1}$ was accepted. Time to completion for this task is displayed in Figure 6(d).
  - *Accuracy.* A statistically significant difference was found when comparing the average scale differences between conditions ($Z = 1.14 \times 10^3, p < 0.05$). Therefore, hypothesis $\mathbf{H_2}$ was rejected for this task, and the alternative hypothesis $\mathbf{AH_2}$ was accepted. Overall accuracy and time to completion per task are displayed in Table 2.
- **Rotate.**
  - *Time to Complete.* No significant differences were found in completion time between the two interaction conditions in 3.1. Therefore, hypothesis $\mathbf{H_1}$ was accepted for this task. Time to completion for this task is displayed in Figure 6(e).
  - *Accuracy.* No significant differences were found when comparing the average angle difference in (x, y, z) between conditions in 3.1. Therefore, hypothesis $\mathbf{H_2}$ was accepted.Overall accuracy and time to completion per task are displayed in Table 2.

4.2   Interaction metrics

- **NASA-TLX.** The metaphoric mapped interaction using the HoloLens v1 condition obtained an average NASA-TLX score of 60.18 (SD = 17.74) while the isomorphic mapped interaction condition using the Meta2 scored a 44.61 (SD = 16.80). The metaphoric mapped interaction paradigm using the HoloLens v1 was perceived to have a significantly higher workload, with significant differences found between conditions ($Z = 15.0, p < 0.05$). Consequently, we rejected $\mathbf{H_3}$ for NASA-TLX metric and accepted the alternative hypothesis that the interaction paradigms within the device conditions had an effect on the perceived workload.
- **System Usability Scale (SUS).** Metaphoric mapped interaction condition using the HoloLens v1 obtained an average SUS score of 55.12 (SD = 21.32) while the isomorphic mapped interaction group using the Meta2 scored a 67.25 (SD = 16.52). Scores can be labelled as just acceptable for the HoloLens v1 condition and 'OK' for the Meta2 interaction group [9]. Significant differences were found between conditions when comparing these scores ($Z = 41.5, p < 0.05$). Consequently, we rejected $\mathbf{H_3}$ for SUS, accepting the alternative hypothesis that the interaction paradigm within the device conditions used had an effect on perceived usability.
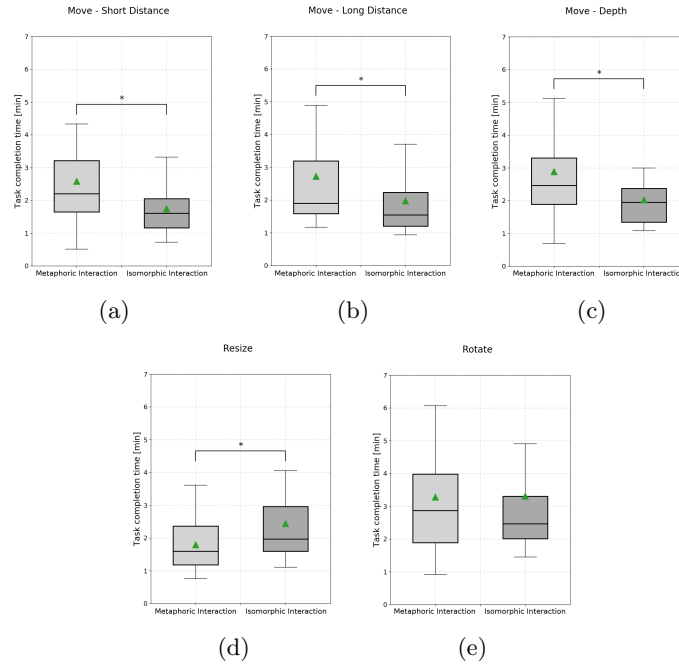
Fig. 6. Completion time per task: 6(a) Move - Short distance task; 6(b) Move - Long distance task; 6(c) Move - Depth task; 6(d) Resize task; 6(e) Rotate task.

– **Post-Questionnaire.** Participants were asked to complete a post-test questionnaire to gain a better understanding of their preferred interaction paradigm. They were asked to rate the performance of each of the devices (Microsoft HoloLens v1 and Meta's Meta2), in terms of their perceived FOV, hand recognition, tracking accuracy and preferred interaction method.

  • **Ease of Use:** Participants were asked to rate the ease of use of both devices using a Likert scale from 1 (Very difficult) to 7 (Very easy). The average score for the HoloLens v1 was 4.2 (SD: 1.75) and for the Meta2 5.18 (SD: 1.37). No significant differences were found between conditions.

  • **Field of View (FOV):** Participants were asked to rate the FOV of each device and if they noted any traking differences. FOV was rated using a Likert scale ranging from 1 (very narrow) to 7 (very wide). The HoloLens v1 scored a 3.45 (SD: 1.35) while the Meta2 received a rating of 4.25 (SD: 1.60). Additionally, 16 participants out of 20 reported on noticing FOV differences between devices; of those 16, 10 agreed that the Meta2 had a wider FOV, which matched the specifications of the device. No significant differences were found between conditions.

  • **Hand tracking accuracy:** Participants were asked to rate the hand tracking accuracy of the device with a Likert scale ranging from 1 (very inaccurate) to 7 (Really accurate). Both devices scored similar results,

Table 2. Overall accuracy and time to completion per task and device interaction paradigm. Time to completion is reported in [min] while accuracy is reported in [cm] for move and resize tasks and in degrees for rotate task. *(* denotes statistical significance between conditions and bold face the best result)*

	Time to Completion		Accuracy	
	Metaphoric	*Isomorphic*	*Metaphoric*	*Isomorphic*
Move-Short	2.58 min*	**1.75 min***	14.01 cm*	**6.86 cm***
	(SD: 1.58 min)	**(SD: 0.73 min)**	(SD: 31.32 cm)	**(SD: 2.18 cm)**
Move-Long	2.72 min*	**1.97 min***	18.85 cm*	**6.57 cm***
	(SD: 1.84 min)	**(SD: 1.05 min)**	(SD: 47.14 cm)	**(SD: 1.48 cm)**
Move-Depth	2.88 min*	**2.03 min***	19.86 cm*	**6.85 cm***
	(SD: 1.70 min)	**(SD: 0.94 min)**	(SD: 39.00 cm)	**(SD: 2.06 cm)**
Resize	**1.79 min***	2.44 min*	**4.12 cm***	13.67 cm*
	(SD: 0.79 min)	(SD: 1.33 min)	**(SD: 9.60 cm)**	(SD: 12.42 cm)
Rotate	**3.28 min**	3.32 min	**28.09°**	33.33°
	(SD: 1.96 min)	(SD: 2.35 min)	**(SD: 41.99°)**	(SD: 51.25°)

with the HoloLens v1 scoring a 4.5 (SD 1.5) out of 7 and the Meta2 a 4.15 (SD: 1.15).18 out of the 20 participants reported on noticing tracking differences between devices; 12 of those stated that the HoloLens v1 had a more stable tracking. No significant differences were found between conditions.

- **Movement:** Despite the Meta2 being a tethered device, none of the participants reported feeling their movement or interaction constrained by this in any way. However, for larger interaction spaces this needs to be considered as a limitation.
- **Preferred interaction method:** When asked to select their preferred general interaction method, 18 participants chose the Meta2 while the remaining 2 selected the HoloLens v1. When asked in task by task basis, participants reported the following:
  * *Move tasks (long, short and depth):* 17 participants (85%) preferred the isomorphic mapped interaction paradigm, while 2 (10%) selected the metaphoric mapped paradigm and 1 (5%) reported no preference.
  * *Resize task:* 11 participants (55%) preferred the isomorphic mapped interaction paradigm, while 7 (35%) selected the metaphoric mapped paradigm and 2 (10%) reported no preference.
  * *Rotate task:* 13 participants (65%) preferred the isomorphic mapped interaction paradigm, while 7 (25%) selected the metaphoric mapped paradigm and 2 (10%) reported no preference.
- **Observations and feedback.** Twelve participants (60%) reported fatigue or pain during the metaphoric mapped interaction condition while 8 (40%) reported it during the isomorphic mapped interaction condition.

- **Metaphoric mapped paradigm** *(using the HoloLens v1)*: Five participants reported arm tiredness and ache during the interaction. One participant reported "hard to control objects in the distance but smoother interaction" while another said "the inability to get closer to the objects using the HoloLens v1 did limit my accuracy in the tasks".
- **Isomorphic mapped interaction paradigm** *(using the Meta2)*: None reported arm fatigue or pain in this condition. One participant reported on this interaction being "fun and intuitive, finding easier with the moving aspect to move objects below me more than above", another said "The interaction felt natural and intuitive. The resize was the most difficult as sometimes the tracking of the hands was less accurate." while a third one reported on "the headset being too heavy and having issues with the tracking".

5   Discussion

This study was designed to compare the interaction paradigms and design form factors beneath the two most current, as of the writing of this paper, commercial wearable AR solutions for freehand interaction, namely the Microsoft's HoloLens v1 and the Meta's Meta2. Therefore we presented a formal study to evaluate naive users' response to these devices in a task-based scenario for object manipulation in AR.

As highlighted by Gabbard *et al.* [21], it is important to understand the fundamental perceptual and cognitive factors for new technologies that alter the way humans perceive the world as in immersive technologies (Virtual Reality (VR), AR and Mixed Reality (MR)). Previous research has found promising possibilities for natural interaction in AR applications [12, 17, 42], and while both interaction paradigms had proven to offer a viable interaction method [29], the usability, interactivity and precision of these two devices for object manipulation has not been fully addressed before. We started by exploring and analysing precision metrics, namely time to completion and accuracy. Considering solely these metrics, a compromise between devices' interaction paradigms and spatial considerations need to be considered to optimise task precision and accuracy. Results suggested that the isomorphic mapped interaction paradigm, with all the design considerations associated, was helpful with tasks that required the movement and position of the virtual objects in specific locations in the space (as with move tasks reported in 3.4). This may be linked to the way we manipulate real world objects, as people move (translate) objects around with their hands, placing them in positions where they are physically optimal for each task [25]. However, the metaphoric mapped paradigm using the HoloLens v1 was significantly more accurate and it took significantly shorter time for the uniform resize manipulation task. This may be linked again to users past experience performing shape scale transformations on desktop environments rather than in reality where this form of object scale manipulation is not commonly feasible. Finally the rotation task presented no significant differences between conditions

showing consistency across both interaction paradigms. These results show the convenience of adopting a mixed methods interaction paradigm design for task manipulation in AR, enabling users to move objects around using an isomorphic spatial interaction while enabling a metaphoric gesture based interaction for transformation tasks (i.e. rotate and resize).

Reviewing interaction metrics, the metaphoric mapped *(using the HoloLens v1)* interaction condition was reported to have a significantly higher task load index than the isomorphic interaction paradigm *(using the Meta2)*; as reported in section 4.2. This trend is maintained for the usability score obtained by the two conditions, where the isomorphic mapped interaction paradigm *(using the Meta2)* scored higher in usability according to the SUS results. Usability studies have been deemed helpful to identify design flaws in application concepts at earlier phases of the development of AR [8].

Participants were asked to report on their generally preferred interaction paradigm, where the isomorphic mapped interaction paradigm *(using the Meta2)* received 18 out of 20 votes. The close link between this interaction and real world object manipulations may have an impact on users feeling more confident, as prior experience is a leading contributor to intuitive use [10]. Considering the device form factors, they did not seem to have a determinant effect on the perceived usability, with no statistical differences reported in the post-test questionnaire.

In summary isomorphic mapped interaction *(using the Meta2)* was considered the most intuitive and natural interaction paradigm, with a nearly unnoticeable learning curve. However, when task accuracy and time to completion is key for successfully completing a task, mixed interaction paradigms need to be considered, enhancing these isomorphic interactions with more abstract gestures for certain transform manipulations, While a combination of paradigms may need to be considered, it is key to evolve current interactions to achieve a standardised approach that are device and technology independent.

5.1   Limitations and generalisation

We explicitly limited our study to the tasks and conditions reported in this paper. The interaction conditions metaphoric mapped *(using the HoloLens v1)* and isomorphic mapped *(using the Meta2)* were selected as they were inherent and recommended, as of the writing of this paper, to the two commercial wearable AR solutions used.

The environment was adapted for each condition in 3.1 to comply with both manufacturers' recommendations. This paper presents a comparison study between both devices and interaction paradigms using their ideal environment guidelines for virtual object placement and interaction. Therefore, the use of interaction paradigms and spatial considerations different from the ones outlined in the guidelines [2, 1] were not included in this study.

This study focused around freehand interaction in Mixed Reality, thus, without the use of controllers or wearable devices. Therefore, other interaction metaphors and devices outside this definition were not considered. The overall preference

for freehand natural interaction versus the use of hand-held controllers need to be further explored.

Studied tasks were limited to the transformation tasks reported by [42] to be the most common transform tasks used in the literature [6, 39].

Participants had no previous substantial experience with the devices used in this study, as reported in section 3.6, as we were interested in evaluating naive users' perspective. However, it is being previously explored that training may lead to further improvements in immersive AR systems, as people learn to coordinate perception and interaction [7]. However, this is out of the scope of the presented study and it was not explored.

6   Conclusion and Future Work

We have conducted a study comparing two different freehand AR HMDs and their interaction paradigms associated, namely metaphoric mapped interaction *(using the HoloLens v1)* and isomorphic mapped interaction *(using the Meta2)* in a task based interaction environment. We followed a within participants study design and we reported on interaction and precision metrics. Our results suggest that each device has specific strengths, specially from a precision point of view, with move deemed more accurate under the isomorphic mapped interaction paradigm *(using the Meta2)* and resize being more accurate under the metaphoric mapped condition *(using the HoloLens v1)*. From an interaction perspective, the isomorphic mapped interaction *(using the Meta2)* was perceived as more natural and usable with a higher usability score and a lower task-load index.

Further work needs to be done analysing described tasks in 3.4, i.e. evaluating non-uniform object manipulation (resize task) or alternative shaped objects, as these may lead to differences in accuracy and usability.

Tasks outside the transform category described by [42] were not explored, it may be interesting to further analyse browsing, editing and menu tasks in [42] as these tasks are commonly presented in recent literature [6, 39].

Our findings have some interesting implications for the design and usage of specific AR devices and interaction paradigms for virtual object manipulation tasks. This study has shown that while isomorphic mapped interaction may feel more natural with participants, care must be taken to ensure that the correct interaction paradigm and device are being used, as for environments where accuracy is a requirement, a combination of both paradigms may be key. We also believe that more powerful tracking and visualization devices (i.e. with increased FOV, higher tracking accuracy, device ergonomics) may lead to novel or improved interaction, enhancing usability, comfort and accessibility of AR experiences in the near future.

References

1. Meta Meta2 User Interface Guidelines, (date last accessed 12th of december 2018), http://devcenter.metavision.com/design/user-interface-guidelines-overview

2. Microsoft Hololens Interaction Design Guidelines, (date last accessed 12th of december 2018), https://developer.microsoft.com/en-us/windows/mixed-reality/category/interaction_design

3. Akçayır, M., Akçayır, G.: Advantages and challenges associated with augmented reality for education: A systematic review of the literature. Educational Research Review **20**, 1–11 (2017)

4. Al-Kalbani, M., Williams, I., Frutos-Pascual, M.: Analysis of medium wrap free-hand virtual object grasping in exocentric mixed reality. In: Mixed and Augmented Reality (ISMAR), 2016 IEEE International Symposium on. pp. 84–93. IEEE (2016)

5. Al-Kalbani, M., Williams, I., Frutos-Pascual, M.: Improving freehand placement for grasping virtual objects via dual view visual feedback in mixed reality. In: Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology. pp. 279–282. ACM (2016)

6. Alce, G., Roszko, M., Edlund, H., Olsson, S., Svedberg, J., Wallergrd, M.: [poster] ar as a user interface for the internet of things comparing three interaction models. In: 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct). pp. 81–86. IEEE (2017)

7. Bach, B., Sicat, R., Beyer, J., Cordeil, M., Pfister, H.: The hologram in my hand: How effective is interactive exploration of 3d visualizations in immersive tangible augmented reality? IEEE Transactions on Visualization & Computer Graphics (1), 1–1 (2018)

8. Bai, Z., Blackwell, A.F.: Analytic review of usability evaluation in ismar. Interacting with Computers **24**(6), 450–460 (2012)

9. Bangor, A., Kortum, P., Miller, J.: Determining what individual sus scores mean: Adding an adjective rating scale. Journal of usability studies **4**(3), 114–123 (2009)

10. Blackler, A., Popovic, V.: Towards intuitive interaction theory (2015)

11. Brooke, J., et al.: Sus-a quick and dirty usability scale. Usability evaluation in industry **189**(194), 4–7 (1996)

12. Buchmann, V., Violich, S., Billinghurst, M., Cockburn, A.: Fingartips: gesture based direct manipulation in augmented reality. In: Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia. pp. 212–221. ACM (2004)

13. Cascales, A., Laguna, I., Perez-Lopez, D., Perona, P., Contero, M.: An experience on natural sciences augmented reality contents for pre-schoolers. In: International Conference on Virtual, Augmented and Mixed Reality. pp. 103–112. Springer (2013)

14. Chen, L., Day, T., Tang, W., John, N.: Recent developments and future challenges in medical mixed reality. In: 2017 IEEE International Symposium on Mixed and Augmented Reality. pp. 123–135. IEEE (2017)

15. Ciodota, M., Lukosch, S., Bank, P., Ouwehand, P.: Towards engaging upper extremity motor dysfunction assessment using augmented reality games. In: (ISMAR-Adjunct), 2017 IEEE International Symposium on Mixed and Augmented Reality. pp. 275–278. IEEE (2017)

16. Diaz, C., Walker, M., Szafir, D.A., Szafir, D.: Designing for depth perceptions in augmented reality. In: 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). pp. 111–122. IEEE (2017)

17. Dingler, T., Funk, M., Alt, F.: Interaction proxemics: Combining physical spaces for seamless gesture interaction. In: Proceedings of the 4th International Symposium on Pervasive Displays. pp. 107–114. ACM (2015)

18. Foronda, C., Alfes, C., Dev, P., Kleinheksel, A., Nelson, D., O'Donnell, J.M., Samosky, J.T.: Virtually nursing: Emerging technologies in nursing education. Nurse Educator **42**(1), 14–17 (2017)
19. Funk, M., Kritzler, M., Michahelles, F.: Holocollab: A shared virtual platform for physical assembly training using spatially-aware head-mounted displays. In: Proceedings of the Seventh International Conference on the Internet of Things. pp. 1–7. ACM (2017)
20. Furlan, R.: The future of augmented reality: Hololens-microsoft's ar headset shines despite rough edges [resources_tools and toys]. IEEE Spectrum **53**(6), 21–21 (2016)
21. Gabbard, J.L., Swan II, J.E.: Usability engineering for augmented reality: Employing user-based studies to inform design. IEEE Transactions on visualization and computer graphics **14**(3), 513–525 (2008)
22. Gabriel, E., Jack, M., Mariangely, I.P., Anastacia, M., Winer, E.: Evaluating the microsoft hololens through an augmented reality assembly application. In: Proceedings of SPIE. pp. 10197 – 10197 – 16
23. Hart, S.G., Staveland, L.E.: Development of nasa-tlx (task load index): Results of empirical and theoretical research. In: Advances in psychology, vol. 52, pp. 139–183. Elsevier (1988)
24. Herron, J.: Augmented reality in medical education and training. Journal of Electronic Resources in Medical Libraries **13**(2), 51–55 (2016)
25. Houde, S.: Iterative design of an interface for easy 3-d direct manipulation. In: Proceedings of the SIGCHI conference on Human factors in computing systems. pp. 135–142. ACM (1992)
26. Jankowski, J., Hachet, M.: Advances in interaction with 3d environments. In: Computer Graphics Forum. vol. 34, pp. 152–190. Wiley Online Library (2015)
27. Jetter, J., Eimecke, J., Rese, A.: Augmented reality tools for industrial applications: What are potential key performance indicators and who benefits? Computers in Human Behavior **87**, 18–33 (2018)
28. Jones, J.A., Swan II, J.E., Singh, G., Kolstad, E., Ellis, S.R.: The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In: Proceedings of the 5th symposium on Applied perception in graphics and visualization. pp. 9–14. ACM (2008)
29. Karam, M., Schraefel, M.C.: A taxonomy of gestures in human computer interactions. Tech. rep. (2005)
30. Kim, K., Billinghurst, M., Bruder, G., Duh, H.B.L., Welch, G.F.: Revisiting trends in augmented reality research: A review of the 2nd decade of ismar (2008–2017). IEEE transactions on visualization and computer graphics **24**(11), 2947–2962 (2018)
31. Kruijff, E., Swan, J.E., Feiner, S.: Perceptual issues in augmented reality revisited. In: Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on. pp. 3–12. IEEE (2010)
32. Lampton, D.R., McDonald, D.P., Singer, M., Bliss, J.P.: Distance estimation in virtual environments. In: Proceedings of the human factors and ergonomics society annual meeting. vol. 39, pp. 1268–1272. SAGE Publications Sage CA: Los Angeles, CA (1995)
33. Lasi, H., Fettke, P., Kemper, H., Feld, T., Hoffmann, M.: Industry 4.0. business & information systems engineering, 6, 239-242 (2014)
34. Lee, M., Billinghurst, M., Baek, W., Green, R., Woo, W.: A usability study of multimodal input in an augmented reality environment. Virtual Reality **17**(4), 293–305 (2013)

35. Lee, S., Lee, B., Cho, J., Jang, C., Kim, J., Lee, B.: Analysis and implementation of hologram lenses for see-through head-mounted display. IEEE Photonics Technology Letters **29**(1), 82–85 (2017)
36. Lin, S., Cheng, H.F., Li, W., Huang, Z., Hui, P., Peylo, C.: Ubii: Physical world interaction through augmented reality. IEEE Transactions on Mobile Computing **16**(3), 872–885 (2017)
37. Macaranas, A., Antle, A.N., Riecke, B.E.: What is intuitive interaction? balancing users performance and satisfaction with natural user interfaces. Interacting with Computers **27**(3), 357–370 (2015)
38. Moser, K.R., Swan, J.E.: Evaluation of hand and stylus based calibration for optical see-through head-mounted displays using leap motion. In: Virtual Reality (VR), 2016 IEEE. pp. 233–234. IEEE (2016)
39. Nguyen, H., Ketchell, S., Engelke, U., Thomas, B., d. Souza, P.: [poster] holobee: Augmented reality based bee drift analysis. In: 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct). pp. 87–92. IEEE (2017)
40. Ni, T., Bowman, D.A., North, C., McMahan, R.P.: Design and evaluation of freehand menu selection interfaces using tilt and pinch gestures. International Journal of Human-Computer Studies **69**(9), 551–562 (2011)
41. Piumsomboon, T., Altimira, D., Kim, H., Clark, A., Lee, G., Billinghurst, M.: Grasp-shell vs gesture-speech: A comparison of direct and indirect natural interaction techniques in augmented reality. In: Mixed and Augmented Reality (ISMAR), 2014 IEEE International Symposium on. pp. 73–82. IEEE (2014)
42. Piumsomboon, T., Clark, A., Billinghurst, M., Cockburn, A.: User-defined gestures for augmented reality. In: CHI'13 Extended Abstracts on Human Factors in Computing Systems. pp. 955–960. ACM (2013)
43. Ren, G., O'Neill, E.: 3d selection with freehand gesture. Computers & Graphics **37**(3), 101–120 (2013)
44. Rolland, J.P., Gibson, W., Ariely, D.: Towards quantifying depth and size perception in virtual environments. Presence: Teleoperators & Virtual Environments **4**(1), 24–49 (1995)
45. Santos, B.S., Cardoso, J., Ferreira, B.Q., Ferreira, C., Dias, P.: Developing 3d freehand gesture-based interaction methods for virtual walkthroughs: Using an iterative approach. In: Handbook of Research on Human-Computer Interfaces, Developments, and Applications, pp. 52–72. IGI Global (2016)
46. Sato, M., Suzuki, S., Ebihara, D., Kato, S., Ishigaki, S.: Pseudo-softness evaluation in grasping a virtual object with a bare hand. In: SIGGRAPH Posters. pp. 40–1 (2016)
47. Shapiro, S.S., Wilk, M.B.: An analysis of variance test for normality (complete samples). Biometrika **52**(3/4), 591–611 (1965)
48. Singh, G., Swan II, J.E., Jones, J.A., Ellis, S.R.: Depth judgment measures and occluding surfaces in near-field augmented reality. In: Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization. pp. 149–156. ACM (2010)
49. Suzuki, S., Suzuki, H., Sato, M.: Grasping a virtual object with a bare hand. In: ACM SIGGRAPH 2014 Posters. p. 51. ACM (2014)
50. Swan, J.E., Jones, A., Kolstad, E., Livingston, M.A., Smallman, H.S.: Egocentric depth judgments in optical, see-through augmented reality. IEEE transactions on visualization and computer graphics **13**(3), 429–442 (2007)
51. Swan, J.E., Singh, G., Ellis, S.R.: Matching and reaching depth judgments with real and augmented reality targets. IEEE Transactions on Visualization and Computer Graphics **21**, 1289–1298 (2015)

52. Wachs, J.P., Kölsch, M., Stern, H., Edan, Y.: Vision-based hand-gesture applications. Communications of the ACM **54**(2), 60–71 (2011)
53. Wilcoxon, F., Wilcox, R.A.: Some rapid approximate statistical procedures. Lederle Laboratories (1964)
54. Yung, R., Khoo-Lattimore, C.: New realities: a systematic literature review on virtual reality and augmented reality in tourism research. Current Issues in Tourism pp. 1–26 (2017)