C-AR-D: Contextual Card Based Information with Augmented Reality

GIZING KHALANDI, Department of Industrial Design, Eindhoven University of Technology, The Netherlands
WESLEY HARTOGS, Department of Industrial Design, Eindhoven University of Technology, The Netherlands
YORN THIJSSEN, Department of Industrial Design, Eindhoven University of Technology, The Netherlands
JULES SINSEL, Department of Industrial Design, Eindhoven University of Technology, The Netherlands

Augmented Reality(AR) technology improves with recent developments in head-mounted holographic displays like the Microsoft HoloLens. Big companies are investing large amounts of money in researching and developing metaverse solutions. It is imaginable that the emergent technology of Augmented Reality becomes one of the dominant technologies of everyday life. In this speculative future, head-mounted displays are small, comfortable, and ingrained enough in everyday life. In this near future, we deem it undesired that you become overstimulated due to an abundance of contextual information and pop-ups. This paper presents C-AR-D: a proposition on interacting with real-time contextual information on physical service cards. The Augmented Reality visualizations are confined to the space of the service card to give the user autonomy to choose which information is relevant. This study aims to define a gesture set for inaction with C-AR-D via a gesture elicitation study.

Additional Key Words and Phrases: Augmented Reality, Mixed Reality, Gestures, speculative future

ACM Reference Format:
Gizing Khalandi, Wesley Hartogs, Yorn Thijssen, and Jules Sinsel. 2018. C-AR-D: Contextual Card Based Information with Augmented Reality. In *DUIET '21 - '22: Designing User Interfaces with Emergent Technologies, Okt-Jan, '21-'22, Eindhoven, NL.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

State of the world... We contest that since the introduction of Augmented- and Mixed Reality (AR and MR) technologies efforts have been made to minimize the borders between the digital and the physical world. These make it possible to create future visions as the Metaverse [14] and Hyperreality [24]. Technologies as AR are already integrating into the everyday life by implementations as navigation, gaming (Pokemon Go), factory work and more [8, 19]. New head mounted displays are introduced on the market, but we see a treat over stimulation of everyday life. Artists and designers are already speculating and warning about this overstimulated future, where everything wants the users attention, for example, in the video named HYPER-REALITY by Keiichi Matsuda [11]. We deem this speculative future is unwanted but see a future where MR Head Mounted Displays are the new norm in modern society as highly likely. A question arises how people will interact with these MR application, how it will be used in different contexts and what will be the place of physical products when the digital and physical lines blurred. Our hypothesis is that in the future where MR Mounted Displays are the norm, physical and digital artifacts are merged, creating new services and interactions between user and product. Therefore, we designed C-AR-D a physical card based Augmented Reality information system where we confine the visualization the the physical dimensions of the service card. This study limits itself to gesture based input for MR in this speculated future, where participants are asked to perform gestures what are defined natural and comfortable. The contribution of this paper is a first exploration into a set gestures for interacting with a hybrid form of Mixed Reality interaction.

2 RELATED WORKS

2.1 Augmented Reality

AR is becoming more popular [22] and many applications across multiple domains are developed on AR hardware [22] such as the Microsoft HoloLens or similar Head-Mounted Displays (HMD) [7]. The latest advances in AR use hand gestures for interaction. Yet, little is currently known about user preference, and behaviour for gestures in AR [17]. Previously done, AR researchers have shown the use of hands as a form of input still requires a lot of research as current knowledge still has deficiencies.

2.2 AR combined with hand gestures

By being strict about all the different possibilities of applying AR to consumer products, the gestures that the systems have to recognize are limited. Although it is of course convenient for users to learn these standardized gestures by heart, making it a habit relatively easy, e.g. zooming in and out on a touch screen, it is for the developers of applications that run on this knowledge limiting in the possibilities that the hardware and software could offer. This is due to the fact that the appropriate gestures that are considered "natural" depend on the situation in which they apply [7]. This is a wrong approach to the problem. It should not be that users learn gestures by memory as a foreign language but that the actions feel intuitive, which also opens new possibilities for the developer.

2.3 Gesture recognition and manipulation

To use gestures for MR, they must be recognized and classified, which is a complex problem. The recognition of gestures is considered an emerging technology, but it has been a research subject for a long time [6]. Mitsubishi provided a method for recognizing hand gestures already in 1995, based on a pattern recognition technique developed by McConnell that uses histograms of local orientation [16]. More current studies on gesture recognition applicable to this study are [26] said research into developing a recognition system for continuous natural gestures, using a system capable of tracking 3D hand gestures. Resulting in a quite accurate system -95.6 percentage accuracy on isolated gesture recognition, and 73 percentage recognition rate on continuous gesture recognition- where they applied it to navigate in Google Earth (3D). The challenges are in the signal accuracy and signal interpretation that work in a real-world environment. The study "Physics-based Hand Interaction with Virtual Objects" [12] examines a more direct way of manipulating and interacting with virtual objects through a self-designed system, with the aim of eliciting more realistic responses with virtual objects. The reaction between the virtual objects and the hands of the participants can be quite different from that between real objects and hands what can be an issue sometimes, therefore, according to the study, further research needs to be done with the application in real-time interaction to improve this emerging technology, so it can be applied in products. Gestures that directly interact with the virtual objects are more difficult to apply and are therefore less commonly used for commercial purposes. These interaction gestures present additional challenges because they require the application to make many interpretations about the hand in the camera's field of view, such as depth, occlusion, size, and movement [13]. Can introducing a tangible element into the interaction solve the problems of recognizing these gestures and improve the interaction?

2.4 Tangible interaction

The application of tangible elements in the AR environment could possibly improve the problems surrounding gesture recognition and interaction between AR elements. The user will have a haptic control, which research shows to be in demand. The HCI studies of [5, 20] show the advantages of using a tangible product as a part of the interaction combined with a different technology, resulting in MR. By combining a physical product with the AR environment, the typical problems mentioned above may be addressed. For this, it is interesting to look at the natural gestures that the participants in the study make during implementation, as no concrete research has been done yet.

2.5 AR in navigation

The study focuses on navigating on a map and interacting with the system. Navigation systems are widely used in different sectors (e.g., automotive, hiking, etc.) using software from companies such as Google Maps. Most of these systems are touch screen based, where the user can move his finger around the map, but there are also systems that already use forms of AR, e.g., automotive industry [10]. The study [21] has investigated the combination of AR in navigating with a digital map for freehand gestures using a horizontal intangible map display. The study compared different ways of moving around in this digital environment with arm gestures by looking through AR displays. The conclusion of the study is that their designed system met the requirements of conventional digital navigation and included the added value of the AR system. No further research has been done into the use of tangible objects as aids to navigation through the digital map or what the user would do out of the natural feeling.

3 METHODS

This section describes the methods and techniques used in this study. First, the design practice will be discussed, followed by the user test and protocol, and finally, the analysis.

3.1 Demonstrator application; C-AR-D

The C-AR-D is a Mixed Reality application in which a 3D map is visualized on a physical card (see figure 1) to show real-time data, which can be used for various services. The service chosen for this study is to guide new students at Eindhoven University of Technology (TU/e) in a speculative future where they would wear AR glasses. However, since AR glasses were not available, this study will be tested through a Wizard-of-Oz methodology. Wearing AR glasses will be faked by using a webcam and screen setup, which will be further explained in section 2.1 Setup. The service chosen is to guide new students of the TU/e through AR glasses over the university campus. The campus map will emerge on top of a student card when they hold the card in front of them while wearing the glass. The student will interact with the map and use different gestures that feel most natural during the experience. The aim of this study is to specify a user-defined, natural feeling set of gestures for in a public context.

3.2 Participants

The targeted user group in this study are starting TU/e students that do not know the campus yet and need help navigating. Therefore, first-year students following a Bachelor's or Master's degree and pre-Master students could validate the study and C-AR-D.

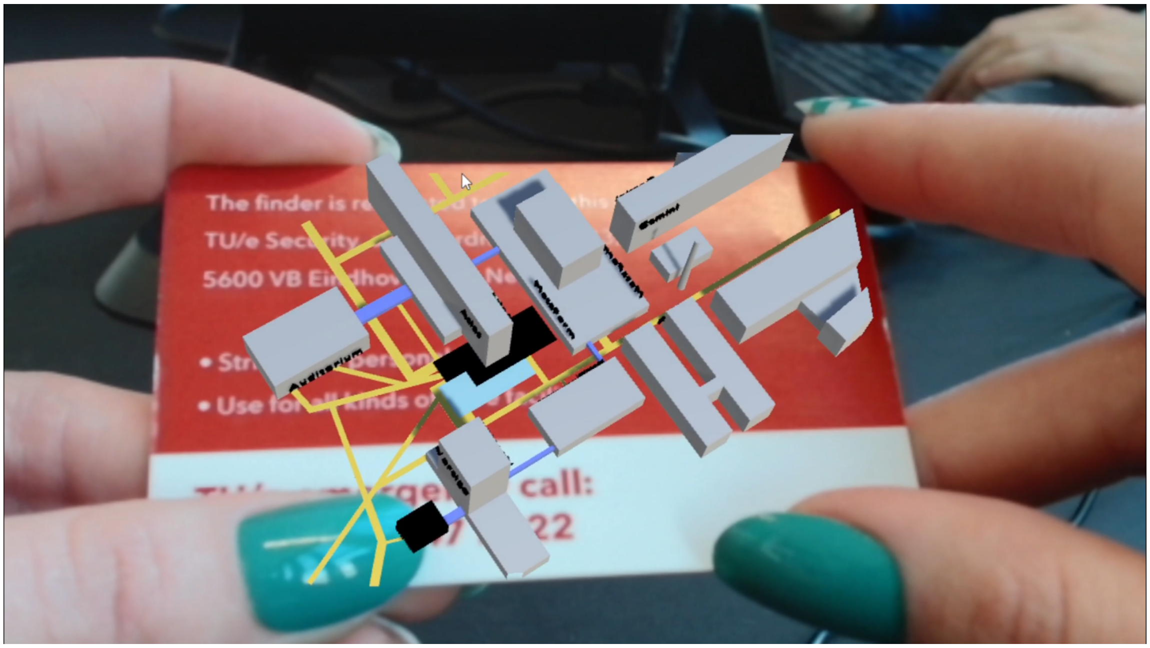

Fig. 1. Demonstrator prototype

3.3 Procedure

As mentioned before, C-AR-D will be tested through a Wizard-of-OZ technique, where the researchers will manipulate the application in such a way that the participant feels as if the interaction is real [4, 25]. he user test will be done following; a webcam will be placed between the participant and a model on a desk (see figure 2). The webcam will record the participant's interaction with the map and use their hands. The webcam footage will be sent to a laptop where the application runs in the game engine Unity [23] and AR output is generated with Vuforia [1]. The researcher manipulates the interaction via a keyboard in Unity, and this footage will be sent to an external screen placed in front of the participant. The participant can use the card and different gestures to fulfill the assignments given.

3.4 User test

The user test starts with the participant filling in the consent form and a short explanation of the project. When they agree, the participant will be seated in front of the screen. During the user test, the footage in Unity will be screen captured, showing the participants' hands and manipulating the researcher. The participant will be asked what gesture they would perform to do the following tasks on the map while following a Think-aloud protocol [9]: moving, rotating, zooming in and out, and finally setting a waypoint. When the participant has decided on a gesture for all tasks, they will be asked to perform each gesture three times, after which they will fill in a small questionnaire about each gesture, one by one. The test ends with a short interview to ask three in-depth questions about the gestures performed (see appendix I). The user study aims to find the most natural gestures to serve in the application.

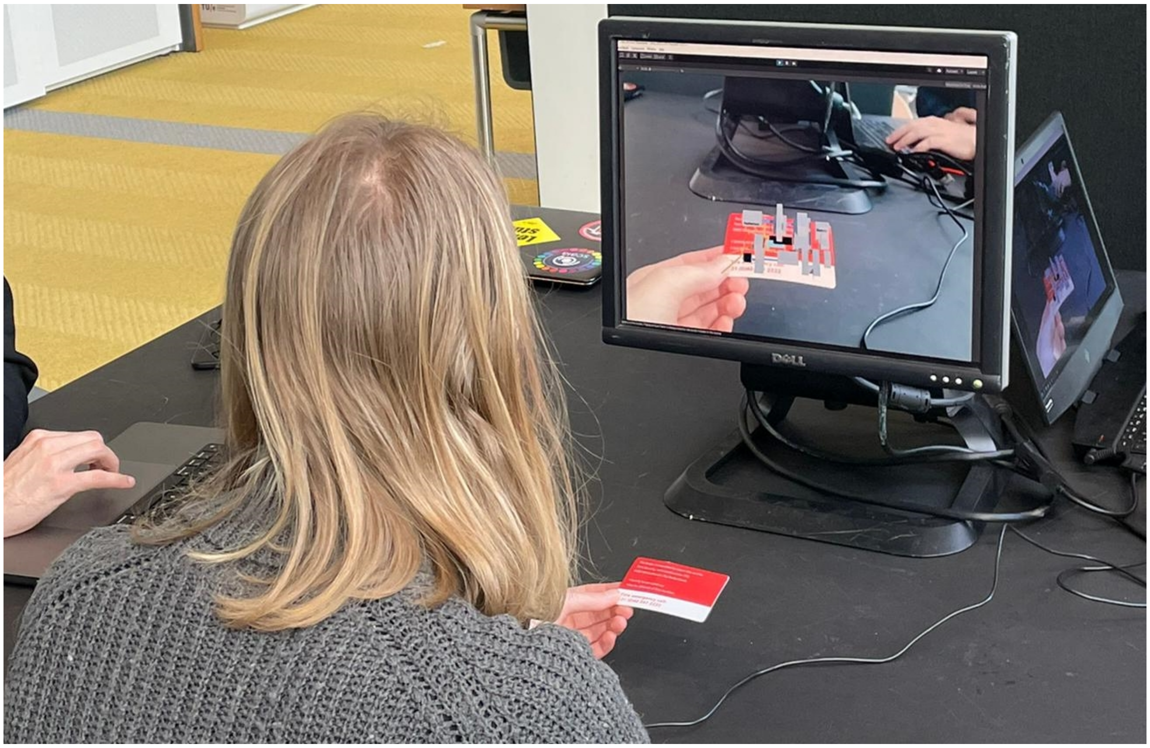

Fig. 2. Participant interacting with the setup.

3.5 Data collection

The data collected during this study is qualitative as quantitative. The qualitative data consists of the screen capture and voice recording during the user test and the recording and notes of the interview. Important quotes will be used to empathize with the different gestures made and why. The quantitative data consists of data from the questionnaire in Microsoft forms that the participant needs to fill in to evaluate the data.

3.6 Analysis

The aim of the analysis is to define which gestures performed during the study are found natural and comfortable doing in public. The qualitative data will be used to analyse the thoughts of the participant during the test and why certain movements are made. The quantitative data will be translated into a variety of diagrams to visualize the results of these data from the Likert scale[3, 15], among which the different gestures made, how easy they are to perform and how comfortable participants find the gesture to perform in public [18].

4 RESULTS

In total six participants participated in the gesture study. Each participant was asked to come up with a gesture for each of in total five different interactions, so in total 30 gestures were collected. A general overview of the results will be given first after which the results of each interaction will be given. This section ends with qualitative insights from the interviews.

4.1 General overview

In total 30 gestures were collected, of which 19 different gestures were given among the five interactions. Appendix II shows the frequency that a gesture was suggested in one of the five interactions. For example, the gesture "swiping with one finger on card" was suggested as being natural for interaction 1: Move three times. For the gesture that was suggested the most for each interaction, the average of each 7-point Likert scale was calculated resulting in general scores for that interaction, as well as average scores for how easy to perform, how comfortable to perform and how comfortable the gesture is to perform in public. If there was not a gesture suggested the most for a given interaction, the scores of the Likert scales of all gestures that were suggested the same number of times were calculated. Based upon these scores, a set of gestures was defined.

4.2 Moving the map

The first gesture that participants had to come up with was for moving the map. Four different gestures were given, amongst one was given three times: "Swiping with one finger on card". This gesture was averagely rated 6.67 on a 7-point Likert scale. This rating was the same (6.67) for the easiness to perform this gesture. The gesture was rated 5.33 on comfortability to perform, as well as comfortability to perform in public.

4.3 Rotate the map

For the second interaction, rotating the map, a total of five different gestures were suggested. "Using thumb and index finger vertically above card" was suggested two times and was averagely rated 6 on the 7-point Likert scale. It was rated 6.5 on easiness and a 6 on comfortability to perform the gesture. However, the 2 participants suggesting this gesture did find it relatively less comfortable to perform in public, rating it averagely a 4.5.

4.4 Zoom in

For the interaction, zooming in, two gestures were suggested equally, both 2 times: "pinch out using thumb and index finger above card" and "pinch out using thumb and index finger on card". Pinching out above the card was generally rated better (6.5 vs 4.5) as well as on easiness (6 vs 5.5) and comfortable to perform (6 vs 4.5). The participants pinching out on the card felt more comfortable doing this in public (4.5 vs 5). However, the similarity in these gestures is pinching out, the only difference being whether it is above or on top of the card. Pinching out only can thus be seen as the most suggested gesture for this interaction, being suggested four times.

4.5 Zoom out

Just like for zooming in, there are very similar suggested gestures for zooming out. All participants suggested gestures as being natural for zooming out that have a form of pinching in. The difference is again in whether participants perform the gesture above or on top of the card, or in which fingers were used. Generally, pinching in is averagely rated 5.67 out of 7. The way pinching in was used the most is "pinch in using the thumb and index finger on card". This gesture was averagely rated a 4.67 out of 7. On easiness and comfortability to perform it was averagely rated a 5.3 and it was rated averagely a 5 to perform the gesture in public.

4.6 Setting waypoint

No gesture was most frequent suggested for the interaction of setting a waypoint. Three gestures were equally suggested two times. Again, there is a similar gesture with the only difference being whether it was performed on top of or above the card. In that regard, it can be stated that tapping using the index finger has been suggested four times. The gesture that was best rated for this interaction is "Tap index finger on card" rated averagely a 5.5 or higher for each rating.

4.7 Qualitative results

When asked if the participants would use the gestures in public, only one said that they would not use it in public. All other participants were positive that they would utilize the chosen gestures in public.

Participants indicated that the use of the card was a good concept that was easy to use. One participant stated that the card itself was uncomfortable to use for a longer period. Others have stated that using a tangible artifact makes the interaction feel pleasant and safe in contrast to interacting in thin air.

If the scenario translates to a real world application, all participants stated that they think it will help them find their way around campus. However, remarks regarding a "you are here" indicator and similarity of buildings were given.

5 DISCUSSION

From the results, we define the following set of gestures for interacting in MR with a 3D map displayed on a service card.

5.1 Defined set of gestures

Move: "Swiping with one finger on card"
Rotate: "Using the thumb and index finger vertically above card
Zoom in: Pinch out
Zoom out: Pinch in
Set waypoint: Tap with index finger

We state that these gestures are also the dominant gestures as seen with touchscreen based devices. We consider them in this case as the best practices.

5.2 Changing COVID measurements

Due to the continuous change of the COVID-19 measures we could not conduct the study with the preferred target group. The actual study was conducted with Master students from the Industrial Design course "Designing User interfaces with Emerging Technologies", who are able to empathize with the scenario while probably knowing the different buildings and locations of the TU/e campus.

5.3 Limitations

The user study was conducted via a webcam in combination with a screen. This is due to the fact that there was no head-mounted holographic display in our reach. When asking the participants if they felt limited in doing the gestures they all answered that they did not feel limited. We are aware that the scenario was tested in a passive way were the participants were sitting inside instead of standing up outside and on the move. Another limitation was the technology that we used. We used Vuforia [1] for the augmented reality and were not able to remove occlusion of the hand between the card and the camera in the visualization. Only one scenario is tested were a map was shown on a card. Other scenarios that utilize other service cards like a bankcard or a public transport card are considered future work.

5.4 Interactions

Most participants used typical gestures that you would see when interacting with a touchscreen. It seems that the physical card was perceived as an application in 3D as you would see on a smartphone. Multiple participants noted the resemblance to Google Maps and the difference that this was in 3D. Translation and manipulation gestures were very similar to touchscreen based gestures, like "drag to move" and "pinch to zoom". We saw a difference in participants that used the gestures in the air in contrast to participants that did the interactions on the card itself. Future studies can research the need of defining that is necessary to make a distinction between the interaction patterns or that it is best practice to program them both. This pattern reminds of Skeumorphism [2] were metaphors can be used as interaction clues on how to interact with an interface. In our study, the interaction patterns used in smartphone interaction were dominant. More research on mixed reality interactions and gestures is needed to verify if the dominant interactions held up in a world were mixed reality is the norm. What we saw is that direct feedback is important for the participants. Because of the direct feedback, some participants forgot that the study was a partial Wizard of Oz and their gestures were not actually recognized. In the setup setting a waypoint gave no direct feedback, this resulted in confusion by the participants. It is worth noting that participants linked zooming in and out as one gesture, when they were zooming in they were already zooming out, we assume this is because of the direct feedback in our setup.

5.5 Physicality of AR

Using a physical card as a restricted form for AR visualization was received well by the participants. It was an anchor for where the interaction needs to happen. The tracking of the student card was not perfect and resulted in a few glitches and misalignment's. Future research could try to fix the tracking issues and make the recognition of the the student card more reliable.

6 CONCLUSION

In this paper, we have presented C-AR-D. A contextual card-based information system that utilizes augmented reality confided to a physical service card. We conducted a gesture elicitation study with the Wizard of Oz methodology on the scenario of a first-year student that needs to find their way on campus. We conclude that gestures based on current best practices of smartphone interactions would fit best to interact with C-AR-D. Our research uses a physical service card as a confined space for AR visualization, and participants received this way of interacting well. Due to the limitations, future work is needed to gain insights into how participants interact with the system when using a head-mounted holographic display. Also, it could expand upon defining more contextual hybrid interactions to give the user more autonomy over the information they want to have visualized in AR concerning their current context.

REFERENCES

[1] [n.d.]. Getting Started | VuforiaLibrary. https://library.vuforia.com/
[2] [n.d.]. SKEUOMORPH | Meaning & Definition for UK English | Lexico.com. https://www.lexico.com/definition/skeuomorph
[3] I Elaine Allen and Christopher A Seaman. 2007. Likert scales and data analyses. *Quality progress* 40, 7 (2007), 64–65.
[4] Nils Dahlbäck, Arne Jönsson, and Lars Ahrenberg. 1993. Wizard of Oz studies—why and how. *Knowledge-based systems* 6, 4 (1993), 258–266.

[5] Daniel Fitzgerald and Hiroshi Ishii. 2018. Mediate: A Spatial Tangible Interface for Mixed Reality. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188472

[6] William T Freeman and Michal Roth. 1995. Orientation histograms for hand gesture recognition. In *International workshop on automatic face and gesture recognition*, Vol. 12. IEEE Computer Society, Washington, DC, 296–301.

[7] Markus Funk, Mareike Kritzler, and Florian Michahelles. 2017. HoloLens is more than air Tap: natural and intuitive interaction with holograms. In *Proceedings of the seventh international conference on the internet of things*. 1–2.

[8] Mohammad S Hasan and Hongnian Yu. 2017. Innovative developments in HCI and future trends. *International Journal of Automation and Computing* 14, 1 (2017), 10–20.

[9] Riitta Jääskeläinen. 2010. Think-aloud protocol. *Handbook of translation studies* 1 (2010), 371–374.

[10] Richie Jose, Gun A. Lee, and Mark Billinghurst. 2016. A Comparative Study of Simulated Augmented Reality Displays for Vehicle Navigation. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction* (Launceston, Tasmania, Australia) *(OzCHI '16)*. Association for Computing Machinery, New York, NY, USA, 40–48. https://doi.org/10.1145/3010915.3010918

[11] Keiichi Matsuda. 2016. HYPER-REALITY. https://www.youtube.com/watch?v=YJg02ivYzSs

[12] Jun-Sik Kim and Jung-Min Park. 2015. Physics-based hand interaction with virtual objects. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 3814–3819.

[13] Byungsung Lee and Junchul Chun. 2009. Manipulation of virtual objects in marker-less AR system by fingertip tracking and hand gesture recognition. In *Proceedings of the 2Nd International Conference on Interaction Sciences: Information Technology, Culture and Human*. 1110–1115.

[14] Lik-Hang Lee, Tristan Braud, Pengyuan Zhou, Lin Wang, Dianlei Xu, Zijun Lin, Abhishek Kumar, Carlos Bermejo, and Pan Hui. 2021. All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda. *arXiv preprint arXiv:2110.05352* (2021).

[15] Rensis Likert. 1932. A technique for the measurement of attitudes. *Archives of psychology* (1932).

[16] Robert K McConnell. 1986. *Method of and apparatus for pattern recognition*. Technical Report.

[17] Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. 2013. User-defined gestures for augmented reality. In *IFIP Conference on Human-Computer Interaction*. Springer, 282–299.

[18] Thammathip Piumsomboon, Adrian Clark, Mark Billinghurst, and Andy Cockburn. 2013. User-Defined Gestures for Augmented Reality. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) *(CHI EA '13)*. Association for Computing Machinery, New York, NY, USA, 955–960. https://doi.org/10.1145/2468356.2468527

[19] Philipp A Rauschnabel, Alexander Rossmann, and M Claudia tom Dieck. 2017. An adoption framework for mobile augmented reality games: The case of Pokémon Go. *Computers in Human Behavior* 76 (2017), 276–286.

[20] Elif Salman, Ceylan Besevli, Tilbe Göksun, Oğuzhan Özcan, and Hakan Urey. 2019. Exploring Projection Based Mixed Reality with Tangibles for Nonsymbolic Preschool Math Education. In *Proceedings of the Thirteenth International Conference on Tangible, Embedded, and Embodied Interaction* (Tempe, Arizona, USA) *(TEI '19)*. Association for Computing Machinery, New York, NY, USA, 205–212. https://doi.org/10.1145/3294109.3300981

[21] Kadek Ananta Satriadi, Barrett Ens, Maxime Cordeil, Bernhard Jenny, Tobias Czauderna, and Wesley Willett. 2019. Augmented reality map navigation with freehand gestures. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 593–603.

[22] Anna Syberfeldt, Oscar Danielsson, and Patrik Gustavsson. 2017. Augmented reality smart glasses in the smart factory: Product evaluation guidelines and review of available products. *Ieee Access* 5 (2017), 9118–9130.

[23] Unity Technologies. [n.d.]. Unity Real-Time Development Platform | 3D, 2D VR & AR Engine. https://unity.com/

[24] Benjamin Woolley. 1993. *Virtual worlds: A journey in hype and hyperreality*. Benjamin Woolley.

[25] Ying Yin and Randall Davis. 2010. Toward Natural Interaction in the Real World: Real-Time Gesture Recognition. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction* (Beijing, China) *(ICMI-MLMI '10)*. Association for Computing Machinery, New York, NY, USA, Article 15, 8 pages. https://doi.org/10.1145/1891903.1891924

[26] Y Yin and DR Toward. 2010. Natural interaction in the real world: realtime gesture recognition. In *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction (ICMI-MLMI'10)*.