ExplorAR - Augmented Reality (AR) Creation Platform

1.0 Abstract

Augmented Reality (AR) is a technology that combines the digital world and real-world assets into one unified experience. This HCI 594 Capstone project aims to study and design an AR platform called ExplorAR that is targeted at users interested in augmented reality. After analyzing 8 existing AR platforms (competitive analysis) and performing a preliminary literature review, our team conducted live user interviews and a hybrid card sort. Using the findings of the interviews and card sort, we proposed a "no-code" platform for all types of users (expert and novice) to get inspired by an easy-to-use process for creating and sharing AR content. We believe our platform, ExplorAR, will be better adapted to the needs of any type of AR user than the existing AR platforms that we reviewed for our competitive analysis.

Our research showed that although people found AR to be exciting and interesting, there is a sense of mistrust and lack of interest among people in implementing AR into their environment. In our competitive analysis of 8 existing AR platforms, we discovered that most solutions lacked clear directions and processes for users to understand the AR platform system, and primarily catered to a specific type of user. Our initial idea of an AR platform focused on content creators specifically who would benefit from adding AR to their content, to be shared with their followers. However, the recruitment for our interview phase did not produce an ample sample size. Therefore, we transitioned to include all users interested in learning and creating content on AR platforms. From our nine user interviews, we identified two user types: novice and expert. Novice users wanted a platform that did not increase cognitive load. Expert users wanted the freedom to break apart and build AR experiences without being limited to the tools used to build an experience. The usability tests that we conducted for our lo-fi and mid-fi prototypes indicated that participants appreciated our AR-platform and found it clean, interactive, creative, and simple to complete the tasks we provided. While we incorporated most of the feedback to our designs from our first round of usability testing, we believe more emphasis on a guided experience along with a high-fidelity prototype would allow for more interaction and would be an ideal future goal for our platform, ExplorAR.

2.0 Introduction

Augmented Reality (AR) is becoming increasingly prevalent in a wide range of fields such as education, instructional training, entertainment, and social media. AR combines aspects of the user's real-world environment with virtual "overlays" that display contextual information in real-time and invite user interactions, (Pietro et al., 2018). While AR has applications for industry, it also appeals to consumers who may use it to play games, try on glasses, or enhance videos on TikTok. Despite its appeal, the ability of the general population to create their own AR experiences is lagging. Most AR creation tools involve expensive professional software and equipment and are geared toward programmers and developers. According to Gasques et al. (2019), many AR tools "still require a high degree of programming proficiency, presenting multiple barriers and a steep learning curve for designers or non-programmers." Leiva et al. (2020) states, "the need for programming knowledge hinders exploration and excludes non-technical designers." We also noticed a "Catch-22" conundrum in that developers may have the programming skill to create an AR application, but they "lack the content knowledge." (Martinez et al., 2014). Conversely, hobbyists, educators, DIYers and influencers may possess vast content knowledge but have little to no coding expertise to translate their content into an AR experience. Our team wondered, what if there was a way to have the best of both worlds? The motivation for our project was to “democratize” the AR authoring process so that anyone, regardless of their ability to code, animate, or create 3D models, could find success in producing an AR experience. We defined our problem statement as, “How might we envision a "no-code” augmented reality creation platform for all?”

When we began exploring the topic, our original aim was to devise an AR creation platform specific to the DIYer who posts on YouTube. The initial focus of our research targeted how 3D augmented reality experiences for “DIY” and assembly tasks are superior to 2D printed instructions (Koutitas et al., 2019). Our literature review revealed that instructional-based AR could provide superior benefits for the following human factors: 1) Reducing cognitive load and errors when performing spatial tasks (Chidambaram et al., 2021 and Herbert et al., 2022); 2) Improving speed and reliability (Chidambaram et al., 2021); 3) Increasing user motivation (Koçak et al., 2019); 4) Maintaining focus and attention; and 5) Improving abilities in conceptualizing complex content (Koçak et al., 2019). The literature influenced our project because it highlighted the tools, processes, and elements we should include in our platform.

The competitive review we conducted included web-based applications that required no additional hardware, wearables, plug-ins, or round trips to third-party applications to utilize. We carried out our analysis based on three factors: 1) Ease of Use; 2) Quality of Use; and 3) Benefit of Use. We documented the relative strengths and weaknesses for each dimension and then synthesized the best features into our platform and made note of those to avoid. Some of the strengths we included were access to “drag and drop” pre-made asset library elements, start-up templates, distilling projects into a discreet number of steps, and integration with other software and hardware. Weaknesses that we sought to avoid included overwhelming interfaces with too many choices, no guidance as to what to do next, lack of integrated help, lack of visual cues, inability to go back to a previous screen or “undo” a prior action, high learning curve for the first-time user, requiring a secondary app to generate a QR code, and lack of precision when positioning elements. We further identified new features such as voice guidance to help navigate users through the authoring process, and streamlined the elements, properties, interactions, and templates to a manageable number.

We believe in having an AR platform that is a simple standalone “all-in-one" application. Our guiding principles included that our platform is inclusive and accessible, giving the user the ability to create content regardless of their education, technical expertise, age, first language or access to advanced technology. Additionally, our team focused on simplicity, and ease of use and envisioned our product to provide onboarding guidance, encourage users to track progress, and minimize the users' cognitive load. We decided to prototype our platform as a web-app because it provides universal access.

7.0 Conclusion

Our project aimed to develop a "no-code" platform for all types of users (expert and novice) to get inspired by an easy-to-use process for creating and sharing AR content. We leveraged several different methods, such as live user interviews, hybrid card sort, and lo-fi and mid-fi prototype evaluations, to help us understand the users' needs, goals, and pain points. Our research revealed that our platform should focus on four main categories: 1) a clear and simple system with easily recognizable elements, 2) intuitive onboarding, 3) robust and uninterrupted performance, and 4) engaging starter templates to foster creativity and support different learning styles.

We incorporated features such as easy navigation to enable users to find and utilize the primary components to create content. Our platform employed design conventions such as a step-by-step progress bar, verbal guidance, and a set of templates to empower users to create content of their choice. We focused our usability testing goals on understanding how our users complete critical tasks related to the designed features. Our lo-fi usability study indicates that we received positive feedback on our prototype. The task failures were minimal and were easily resolved by simplifying the language and the placements of correct icons.

Our second round of usability testing produced excellent results, with an overall 95% success rate for our tasks. Based on our work so far, we believe our platform would be successful in the market as it was well-received by most participants. Further iterations and testing of our platform would help us implement a more refined user experience for our targeted users.

8.0 References

Barnum, Carol M. (2011). Chapter 7 Conducting a usability test. *Usability Testing Essentials* (pp. 249-285). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-816942-1.00007-1.

Chidambaram, S., Huang, H., He, F., Qian, X., Villanueva, A., Redick, T., Stuerzlinger, W., & Ramani K. (2021). ProcessAR: An augmented reality-based tool to create in-situ procedural 2D/3D AR Instructions. *Designing Interactive Systems Conference 2021,* 234-249. https://doi.org/10.1145/3461778.3462126

Gasques, Danilo et al. "What You Sketch Is What You Get." Extended Abstracts of the 2019 *CHI Conference on Human Factors in Computing Systems.* ACM. 1–6. Web.

Gasques, .R., Jain, A., Rick, S., Shangley, L., Suresh, P., & Weibel, N. (2017). Exploring Mixed Reality in Specialized Surgical Environments. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2591–2598. https://doi.org/10.1145/3027063.3053273

Herbert, B., Wigley, G., Ens, B., & Billinghurst, M. (2022). Cognitive load considerations for Augmented Reality in network security training. *Computers & Graphics, 102*, 566–591. https://doi.org/10.1016/j.cag.2021.09.001

Koçak, Ö., Yılmaz, R.,M., Küçük, S., & Göktaş, Y. (2019). The Educational Potential of Augmented Reality Technology: Experiences of Instructional Designers and Practitioners. *Journal of Education and Future, 15*, 17–36. https://doi.org/10.30786/jef.396286

Koutitas, G., Smith, K.S., Lawrence, G., Metsis, V., Stamper, C., Trahan, M. & Lehr, T. (2019). A Virtual and Augmented Reality Platform for the Training of First Responders of the Ambulance Bus. In *Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '19),* 299–302. https://doi.org/10.1145/3316782.3321542

Leiva, Germán et al. "Pronto: Rapid Augmented Reality Video Prototyping Using Sketches and Enaction." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM. 1–13. https://doi.org/10.1145/3313831.3376160 .

Martínez, H., Laukkanen, S., & Mattila, J. (2014). A new flexible augmented reality platform for development of maintenance and educational applications. *International Journal of Virtual Worlds and Human Computer Interaction*, *2*(1), 18-27.

http://dx.doi.org/10.11159/vwhci.2014.003
Pietro, C., Chicchi, G. I. A., Alcañiz, R.M., . Giuseppe, M. (2018). The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of Literature. *Frontiers in Psychology, 9,* 2086–2086. https://doi.org/10.3389/fpsyg.2018.02086
Zhao, J. Parry, C. J., Anjos, R.D., Anslow, C. & Rhee, T. (2020). Voice Interaction for Augmented Reality Navigation Interfaces with Natural Language Understanding. In 3*5th International Conference on Image and Vision Computing New Zealand* 1-6. https://doi.org/10.1109/IVCNZ51579.2020.9290643.

9.0 Appendices: Supporting Material

9.1 Literature Review

Chidambaram, S., Huang, H., He, F., Qian, X., Villanueva, A., Redick, T., Stuerzlinger, W., & Ramani K. (2021). ProcessAR: An augmented reality-based tool to create in-situ procedural 2D/3D AR Instructions. *Designing Interactive Systems Conference 2021,* 234-249. https://doi.org/10.1145/3461778.3462126

High-level summary: The study tested how accurate and time-consuming 2D training and 3D training were. They found that 3D training videos were more accurate, and the participants preferred it. AR also took significantly less completion time. Making AR training material was much faster. The AR tool is called ProcessAR - instead of adding AR to an existing video, the tool records while an AR performs spatial tasks. Participants used Oculus to record the 3D video. For real-time detection, they used YOLO for object detection. Users do not need any formal knowledge about AR other than a one-on-one session to go over how to use the Oculus and get aquatinted with holding objects in VR.
How does this relate: Since our group wants to explore how to make a platform for content creators, there are ideas in the article that can help to guide us during confusing points since we are all new to AR. Since the article also designed an AR tool related to instructions, there might be ideas we can draw upon for our testing phase.

Fry, C. & Lieberman, H. (2020). Great UI Can Promote the “Do Everything Ourselves” Economy. *In HCI International 2020 - Late Breaking Papers: User Experience Design and Case Studies,* 98–107. https://doi.org/10.1007/978-3-030-60114-0_6

High-level summary: The authors state that UI is now enabling a “Do It Yourself” and “Personal Manufacturing” culture where users can perform their own home repairs even using 3D printed tools. This allows a more sustainable economy and can decrease costs by increasing the longevity of appliances and other individual property.
How does this relate: Although this source was not specific to AR, it does corroborate our hypothesis that the users can benefit from accessing AR instructional material available from content providers to perform “DIY” projects or tasks.

Gasques, .R., Jain, A., Rick, S., Shangley, L., Suresh, P., & Weibel, N. (2017). Exploring Mixed Reality in Specialized Surgical Environments. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2591–2598. https://doi.org/10.1145/3027063.3053273

High-level summary: Rodrigues et al. (2017) examine mixed reality in health care and shed

light on opportunities in advanced surgery. Leveraging Microsoft HoloLens, the team developed a prototype mannequin to add mixed reality capabilities to train on surgeries and a prototype to teach human anatomy.
How does this relate: Understanding how the team built and tested such mixed reality prototypes can guide our team to understand our product's good testing and prototyping practices.

Herbert, B., Wigley, G., Ens, B., & Billinghurst, M. (2022). Cognitive load considerations for Augmented Reality in network security training. *Computers & Graphics, 102*, 566–591. https://doi.org/10.1016/j.cag.2021.09.001
High-level summary: The researchers present a network cabling tutoring system which employs a combination of AR, an intelligent tutoring system (ITS) and a "fading mechanism" which minimizes a user's cognitive load by gradually reducing and finally omitting the number of instructional steps (examples) over time. The authors claim that the Cognitive Load Theory (CLT) supports their research because they show how novice users who gain increasing levels of proficiency with executing a training task require fewer examples because the extra information puts a strain on their cognitive load. The AR tutorial shown to participants first displays "AR Arrows", text-based instruction and checklists that guide the users through the training. If a user makes a mistake, the Intelligent Tutoring System (ITS) identifies where the mistake was made and provides immediate feedback using text, AR and text-to-speech. The results show that the participants made fewer errors when those elements were removed and replaced with summaries and then removed completely in the final task. The authors also present several Cognitive Load Theory effects which influenced their findings: 1) Worked Example Effect (shows step-by-step procedures and instructions of a completed example); 2) Split-Attention Effect is splitting attention between a paper-based manual and the equipment being repaired. The implications for design of this effect is to integrate information in a single UI versus spreading it over several UIs; 3) Modality Effect where audio consisting of short text phrases can be added to supplement the visual information since people do not like to read.; 4) Expertise Reversal Effect: AR may be less effective as user's acquire expertise.; 5) Guidance Fading Effect: Progressively "fade out" worked examples because they require a high cognitive load thereby impacting working memory. ; 6) Human Movement Effect: Moving around the item you are fixing can reduce cognitive load.
How does this relate: The research contains useful information regarding how AR interfaces can adapt based on the user's level of expertise and consideration for cognitive load effects such as including short audio instructions, arrows, and checklists.

Koçak, Ö., Yılmaz, R.,M., Küçük, S., & Göktaş, Y. (2019). The Educational Potential of Augmented Reality Technology: Experiences of Instructional Designers and Practitioners. *Journal of Education and Future, 15*, 17–36. https://doi.org/10.30786/jef.396286
High-level summary: The researchers explored the educational benefits of AR in the fields of science education, social studies, and health. Participants in the study noted that AR could improve retention, attract and maintain user's attention, increase motivation, and improve conceptualization of complex subject matter. It can reduce cognitive load and educational costs. The researchers developed guidelines for developing AR instructional material that include making it realistic and interesting, designing the markers to access the AR, including a variety of multi-media elements, and making it easy to use by providing the appropriate level of guidance.
How does this relate: The researchers include helpful feature considerations and

guidelines that our platform can incorporate which they claim help the user maintain focus and attention, have improved conceptualization of complex information, increase motivation and other cognitive benefits.

Koutitas, G., Smith, K.S., Lawrence, G., Metsis, V., Stamper, C., Trahan, M. & Lehr, T. (2019). A Virtual and Augmented Reality Platform for the Training of First Responders of the Ambulance Bus. *In Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '19),* 299–302. https://doi.org/10.1145/3316782.3321542

High-level summary: EMS participants tested out a 3D recreation of an ambulance bus that is used during mass casualty emergencies. They used an AR system to explore what's called an AmBus, which is a bus that is used in large-scale emergencies, but the training on how to utilize this bus is given with a PowerPoint presentation and an hour long lecture, once a year. The AmBus is also limited in quantity, so most EMS personnel do not get a chance to explore one of these buses. So, the article did a rendition of the AmBus and had participants explore it using Oculus. They also gave the participant timed tasks to see how well they remembered where specific items in the AmBus were. This type of training helped EMS personnel to fully train themselves on how to use the AmBus during emergencies.

How does this relate: This article is another example of how training in AR was a success. The article is especially useful in regards to how the AR experience provided hints and labels on items. Another finding is that participants were able to retain what they experienced in AR much better than paper-based. This article shows the importance of why an AR platform for training/DIY can be useful for content creators.

Lehman, S., Alrumayh, A. S., Kolhe, K., Ling, H., & Tan, C.C. (2022). Hidden in Plain Sight: Exploring Privacy Risks of Mobile Augmented Reality Applications. ACM Transactions on Privacy and Security. *https://doi.org/10.1145/3524020*

High-level summary: In this article, Lehman et al. (2022) investigate the danger and threats that lurk in AR applications where publishers run operations secretly that endanger users' privacy. They not only identify these risks but provide steps to help mitigate these risks from occurring.

How does this relate: When building our VR system, the teams need to ensure that we do not let our proposed solution only works when requested. Providing full transparency on system utilization and data collection requirements to our content creators and users will build trust.

Martin, D., Malpica, S., Gutierrez, D. Masia, B. & Serrano, A. (2021). Multimodality in VR: A survey*. ACM Computing Surveys. https://doi.org/10.1145/3508361*

High-level summary: In this article, Martin et al. (2022) shares the results of a survey that explore the multimodality in VR and the advantages to user experiences, performance and knowledge transfer. The paper further provides a list of critical outcomes that a VR experience can achieve by adding various modalities.

How does this relate: Building a compelling VR product will require the team to think through the various outcomes we would like to achieve, and leveraging the learnings of this paper, we will be able to incorporate specific modalities to our product to enhance the overall user experience. The limitation of adding multiple modalities to a VR system is the increased system requirements to process and output such experiences.

Pietro, C., Chicchi, G. I. A., Alcañiz, R.M., . Giuseppe, M. (2018). The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of Literature. *Frontiers in Psychology, 9,* 2086–2086. https://doi.org/10.3389/fpsyg.2018.02086

High-level summary: This paper looks at the history and evolution of virtual and augmented reality. An AR system should: (1) combine real and virtual objects in a real environment; (2) run interactively and in real-time; (3) register real and virtual objects with each other. Furthermore, even if the AR experiences could seem different from VRs, the quality of AR experience could be considered similarly. Additionally, the authors mention that an AR system must also include a camera able to track the user movement for merging the virtual objects, and a visual display, like glasses through that the user can see the virtual objects overlaying to the physical world. To date, two-display systems exist, a video see-through (VST) and an optical see-though (OST) AR systems. The main difference between the two systems is the latency: an OST system could require more time to display the virtual objects than a VST system, generating a time lag between user's action and performance and the detection of them by the system.

How does this relate: This paper will help us in understanding how the AR technologies have evolved and what the future entails for them. It will allow us to design our application accordingly and also enable us to look at the different perspectives.

Radu, I. (2014). Augmented reality in education: a meta-review and cross-media analysis. *Personal and Ubiquitous Computing, 18*(6), 1533–1543. https://doi.org/10.1007/s00779-013-0747-y

High-level summary: Radu (2014) explores the role of AR in education by examining publications and student learning outcomes in AR immersed learning versus non-AR teaching.

How does this relate: This article will help us in looking at how we can create a platform for teaching the audience how and since this is a meta-review it will help us identify opportunities and what is important to consider in our designs.

Rafael, R., Manoela, D., Daniel, F., Yvonne, L., Vinicius, S., Cristiano, A., Marcelo, T., Veronica, T. (2016). Voxar puzzle motion: an innovative AR application proposed using design techniques. *2016 IEEE Virtual Reality Workshop on K-12 Embodied Learning through Virtual & Augmented Reality (KELVAR),* 11–16. https://doi.org/10.1109/KELVAR.2016.7563676

High-level summary: This work presents and discusses the creation of an application to the Voxar Puzzle, an AR platform that has been built to promote children's education in a fun and lively way. The application was conceived using a well-structured methodology named Create Innovation Pack which is based on the design thinking theory. This method aims at helping to exploit the potential of a technology to deliver value to the user of a field, which in this case is children's education. It applies the use of interdisciplinary teams with complementary skills, in which the technology developers are key members. The goal is to conceive solutions that are both innovative and with high impact potential.

How does this relate: Since we are working on creating a platform for content creators that would allow them to teach others how to complete a DIY project, this article will help us in coming up with a framework on how to approach our solution and also look at what difficulties we might face when developing an AR platform.

Yan, Y., Yi, X., Yu, C. & Yi, S. (2019). Gesture-based target acquisition in virtual and augmented reality. *Virtual Reality & Intelligent Hardware, 1*(3), 276–289. https://doi.org/10.3724/SP.J.2096-5796.2019.0007

High-level summary: This article explores the use of gesture-based approaches for target

acquisition in virtual and augmented reality. The researchers note that a gesture-based interaction happens when a user intends to acquire a target, they use their hands, head, or other physical body gestures that the designed device recognizes and infers the most possible target. In the study, the researchers two main parts of the process: the mental model and the behavior model. They applied a participatory design process in two cases: (1) eliciting grasping gestures of objects for object retrieval tasks in VR; (2) eliciting head movement-based gestures to support hands-free control of AR devices.
How does this relate: When it comes to building an AR platform, it is important for us to understand what gestures we would like to use to perform certain actions within our application. The learnings of this paper will help us in looking at what gestures are appropriate for our platform and how we can incorporate effectively.

Zhao, J. Parry, C. J., Anjos, R.D., Anslow, C. & Rhee, T. (2020). Voice Interaction for Augmented Reality Navigation Interfaces with Natural Language Understanding. In 3*5th International Conference on Image and Vision Computing New Zealand* 1-6. 10.1109/IVCNZ51579.2020.9290643.
High-level summary: This paper introduces VOARLA, which is a natural language understanding (NLU) AR voice interface and helps a courier driver to navigate their way to deliver a package. The researcher of the paper conducts a user study to evaluate VOARLA against an AR voice interface without NLU to investigate the effectiveness of NLU in the navigation interface in AR. They defined three factors to evaluate the two interfaces: accuracy, productivity, and commands learning curve. The results of the study suggested that using NLU in AR increases the accuracy of the interface by 15%. However, higher accuracy did not correlate to an increase in productivity. Additionally, NLU helped users remember the commands on the first run when they were unfamiliar with the system. This suggests that using NLU in an AR hands-free application can make the learning curve easier for new users.
How does this relate: This paper will help us in forming an understanding of how we can integrate verbal guidance in our platform and how it has been used in the different areas so far. It will allow us to design the verbal guidance feature accordingly and also help us to look at the different perspectives.