EuroXR 2023

Proceedings of the 20th EuroXR International Conference

Kaj Helin | Frédéric Noël | Wolfgang Schäfer (Eds.)

VTT TECHNOLOGY 422

VTT TECHNOLOGY 422

EuroXR 2023

Proceedings of the 20th EuroXR International Conference

Kaj Hélin (ed.)

VTT Technical Research Centre of Finland Ltd, Finland

Frédéric Noël (ed.)

Grenoble-INP UGA Génie Industriel / Laboratory G-SCOP, France

Wolfgang Schäfer (ed.)

Zurich University of Applied Sciences, Switzerland

Graphical Design and Technical Editing: Päivi Vahala-Partanen

VTT Technical Research Centre of Finland Ltd, Finland

ISBN 978-951-38-8785-8

VTT Technology 422

ISSN-L 2242-1211
ISSN 2242-122X (Online)
DOI: 10.32040/2242-122X.2023.T422

JULKAISIJA – PUBLISHER

VTT
PL 1000
02044 VTT
Puh. 020 722 111
https://www.vtt.fi

VTT
P.O. Box 1000
FI-02044 VTT, Finland
Tel. +358 20 722 111
https://www.vttresearch.com

Cover image: Rotterdaminfo

EuroXR 2023: Rotterdam, The Netherlands

The focus of EuroXR 2023 conference is on novel VR/AR/MR technologies, including software systems, display technology, interaction devices, and applications. Besides papers on the latest scientific results and highlights from many application fields, the EuroXR conference series aims at creating a unique human-dimension framework, interconnecting European and international XR communities, for knowledge cross-fertilisations between researchers, technology providers, and end-users.

29 November to 1 December 2023

Rotterdam, The Netherlands

Conference organizers

Preface

We are pleased to present these conference proceedings in the VTT Technology series, which contains the papers accepted for the Application, Poster and Demo Tracks of EuroXR 2023, the 20th annual EuroXR conference, being hosted by the Immersive Tech Week from November 29th to December 1st 2023, at De Doelen, in Rotterdam, The Netherlands. This publication is thus a collection of the application papers (talks, posters and demonstrations) presented at the conference. It provides an interesting perspective into current and future applications of VR/AR/MR.

In previous years, under the name EuroVR, the conference has been held in Bremen (2014), Lecco (2015), Athens (2016), Laval (2017), London (2018), Tallinn (2019) and Valencia (2020); and as EuroXR in Milan (2021), Stuttgart (2022) and Rotterdam (2023). The focus of the EuroXR conferences is to present, each year, novel Virtual Reality (VR), Mixed Reality (MR) and Augmented Reality (AR) technologies, including software systems, display technologies, interaction devices, and applications, to foster engagement between industry, academia, and the public sector, and to promote the development and deployment of VR/MR/AR technologies in new, emerging, and existing fields. This annual event of the EuroXR association (https://www.euroxr-association.org/) provides a unique platform for exchange between researchers, technology providers, and end users around commercial or research applications.

We would like to warmly thank the industrial committee chairs for their great support and commitment to the conference, and special thanks go to the local organizing committee for their great effort in making this event happen ☺.

On behalf of the organising committee,

Kaj Hélin

Kaj Hélin, EuroXR AB member, Principal Scientist at VTT Technical Research Centre of Finland Ltd, Finland

Frédéric Noël

Frédéric Noël, Professor at Grenoble-INP UGA Génie Industriel / Laboratory G-SCOP, Grenoble, France

Wolfgang Schäfer

Wolfgang Schäfer, Head AGM Research Hub at the Zurich University of Applied Sciences, Switzerland

Application Program Chairs

- Frederic Noel (Grenoble-INP, Université Grenoble Alpes, France)
- Kaj Helin (VTT, Finland)
- Armin Grasnick (IU Internationale Hochschule, Germany)
- Mario Lorenz (Chemnitz University of Technology, Germany)
- Despina Michael-Grigoriou (Cyprus University of Technology, Cyprus)

Posters and Demos Chairs

- Wolfgang Schäfer (ZHAW Zurich University of Applied Sciences, Switzerland)
- Arcadio Reyes Lecuona (Universidad de Málaga, Spain)
- Hugo Falgarone (CEO & Founder of SkyReal, France)
- Matthieu Poyade (Glasgow School of Art, UK)

General Chairs

- Greg Welch (University Central Florida, USA)
- Patrick Bourdot (Université Paris-Saclay, CNRS, France)
- Benjamin de Witt (Director Immersive Tech Week, The Netherlands)

Scientific Program Chairs

- Krzysztof Walczak (Poznan University of Economics and Business, Poland)
- Gabriel Zachmann (University of Bremen, Germany)
- Omar A. Niamut (TNO, The Netherlands)
- Mariano Alcañiz (LabLENI, Universitat Politècnica de València, Spain)
- Kyle Johnsen (University of Georgia, USA)
- Wolfgang Stuerzlinger (Simon Fraser University, Canada)

Organizing Team

- Euro XR – Mariano Alcañiz-Raya, Patrick Bourdot, Frédéric Noël, Jerome Perret, Arcadio Reyes-Lecuona, Krzysztof Walczak, and Gabriel Zachmann
- Immersive Tech Week – Benjamin de Wit, Rayyan Farrukh, Pierre-Stuart Rostain, Manuel Toledo, and Regina van Tongeren

Table of contents

Demos

Application track

Star Catcher: A Mobile App Using Augmented Reality and Gamification to Engage Users in Art Exhibitions

Yoren Gaffary, and Lisa Brel

[1] *Jeolis Solutions, 12 Cours Sablon, 63000 Clermont-Ferrand, France*
Corresponding author: yoren.gaffary@lojelis.com

Keywords: Augmented Reality, Gamification, Engagement, Mobile App, Art, Museum

Introduction

Why would young people want to go to museums while they already have access to all the information they want online? Today's youth is growing among special effects and video games, and the traditional art industry struggles to attract them in its walls. For instance, the Palais des Beaux Arts museum located in Lille, France, recently proposed to its visitors a temporary exhibit called Open Museum Video Game, with pieces of the exhibit spread across its art galleries. Making things worse, the Covid-19 pandemic stroked, and museums had to limit their number of visitors, or worse, were forced to close their doors for a while.

Recent work highlighted the potential of augmented reality (AR) technology to increase user engagement, for instance in the retail industry (McLean & Wilson, 2019). According to them, interactivity, vividness and novelty are key elements to provide a great user experience. More recently, AR solutions were proposed to give users more engaging museum experiences, tending to increase both user engagement and learning (Gong et al., 2022). We propose to combine AR and gamification approaches, a fairly new combination despite both showed benefits for user engagement and task completion time (Nquen & Meixner, 2019). We aim to improve user experience and engagement toward art exhibitions without altering traditional art. Our solution deploys easily in any ongoing or future art exhibition without the need to alter the exhibition itself.

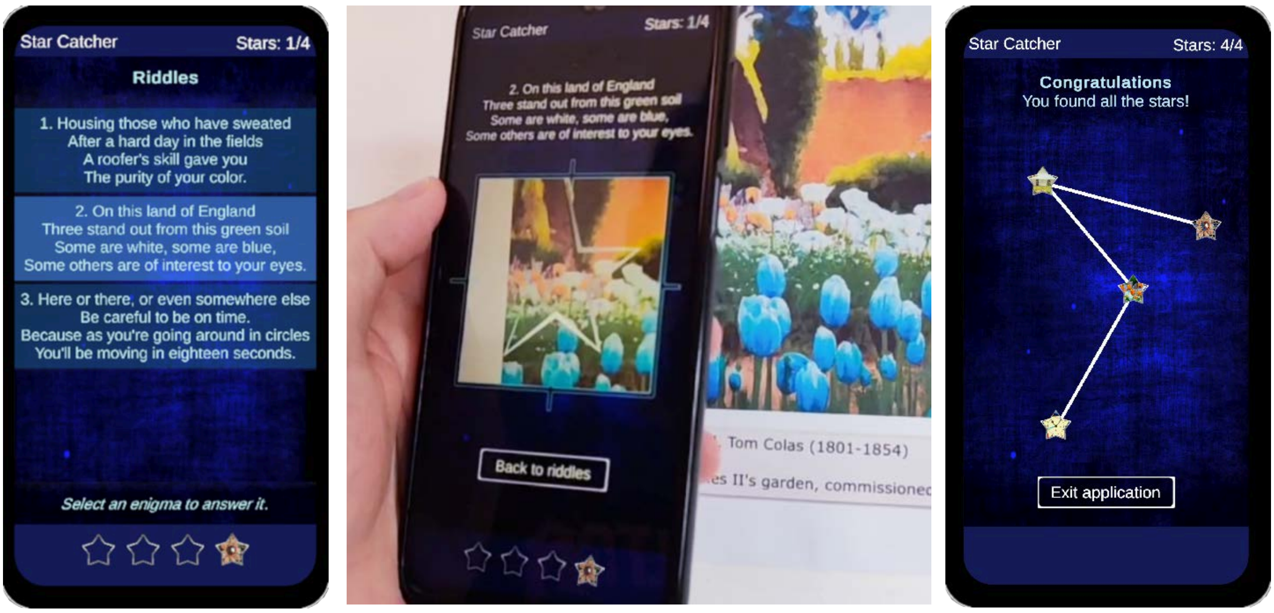
This work presents the development and evaluation of Star Catcher, a mobile app concept for increasing recall and user engagement toward art exhibitions. The app makes the exposed artworks pieces of a hunting game, where users have to solve riddles to find stars hidden in artwork details. While most mobile games as Pokemon Go already use AR, our app use game-inspired mechanisms to engage users in going to museums, and make them learn details about the exposed artworks. This abstract introduces the concept and functioning of the app. It presents a user study evaluating the app in the context of a simulated art exhibition. Conclusion summarizes our concept, discusses our results and exposes in-depth planned future works.

Star Catcher

Star Catcher is a French mobile app designed for art exhibitions. Its main objective is to improve public engagement toward them by allowing people to interact with exposed artworks. The app presents itself as a hunting game, where the player has to find and collect stars that are hidden in exhibited artworks. Riddles associate each star to find with a detail of one of the artworks. For instance, the answer to the riddle "*On this land of England, three stand out from this green soil. Some are white, some are blue, some others are of interest to your eyes*" would be the orange tulips in the background of the picture "Tulip field" (see Figure 1).

The application includes four main screens. The first screen corresponds to a tutorial explaining to the user the goal of the game and how the app works. The second screen displays the different riddles to solve (see Figure 1.a). When the user thinks he or she found the solution to one of the riddles, he or she tap on it to open a new screen displaying on their phone the image issued from the rear camera. The user then scans the detail of the image he or she thinks is the solution to the riddle. If the user is right, a star will be gradually drawn using AR on top of the target detail (see Figure 1.b). The star would always appear to the user co-located with the scanned detail of the artwork, even if he or she moves his or her phone.

Two bars located at the top and bottom of the screen remains on each screen (see Figure 1). The top one displays the number of stars already caught and stars to catch in total. The bottom one displays the collected stars, i.e., the details of artworks related to the riddles they already solved. When a user has collected all the stars, he or she is rewarded with an animation of the stars on the bottom, moving across the phone's screen to draw a constellation (see Figure 1.c).



a. Riddle 2 selection b. Scan of the solution (orange tulips) c. End screen

Figure 1. Applications screenshots of riddles and end screen used for our user study (translated from French), and display of a star drawn using AR over the detail of the picture corresponding to riddle 2.

User Study

This user study aimed to evaluate how Star Catcher improves user experience, and influences memorization of the presented pictures and their description.

Apparatus

This experiment required to simulate an art exhibition to evaluate our app. We used Dall-E 2[1], an AI system able to create original pictures from a textual description, to create six artworks. Each artwork was associated to a fictive title, artist's name and description. This ensures participants could not have seen them prior to the experiment as we aim to evaluate recall.

Participants

30 french-speaking volunteers (14 females, aged from 21 to 56, M=32.3, SD=10.2) were involved in this study. They were randomly assigned to either the experimental or control. Thus, half of the participants used Star Catcher during the tour visit (Application condition, C_A). The app was loaded with riddles related to the exposed artworks. The other half made the tour without Star Catcher, as if they were visiting a classical art exhibition (Control condition, C_C).

Materials

We used the User Experience Questionnaire (UEQ) to evaluate visit experience (Laugwitz et al., 2008). The UEQ consists of 26 items evaluated on bipolar 7-items Likert scales. It evaluates various dimensions of experience, as perceived attractiveness and novelty. We also designed a short questionnaire to collect participants' opinions on the visit through 7-items Likert scales (e.g "*Would you recommend this visit?*"). Participants' memorization of eighteen items was evaluated through MCQ, 6 artworks × 3 questions for each one: a picture detail related to a riddle of the app (e.g. "*What color is the barn's roof?*"), a detail not related to a riddle (e.g. "*How many farmers are represented*?") and an element from picture description (e.g. "*What century is described?*").

Procedure

Firstly, participants were informed we were conducting an experiment to simulate and analyze a museum visit experience. We neither asked them to memorize pictures, nor told them they would have to pass a MCQ after the visit.

Then, participants were left alone in the experiment. No time limit was set for the visit. Six artworks with their description were exposed. The participants were told neither about the real aim of the experiment nor that we used AI-generated pictures until the end of the experiment.

After participants finished the visit, they were seated in another room to ensure they could no longer see the artworks. They first fulfilled a questionnaire concerning their visit experience. They next fulfilled a MCQ evaluating how much information they memorized about artworks and descriptions. At this step we explained to them that we were not aiming to evaluate their personal memory but more globally what people tend to remember to avoid stressing them.

Results

Visit experience. Mann-Whitney test displayed significantly higher scores for items of UEQ in C_A than in C_C (p-value≪0.01). The application made the visit experience feel more original and innovative, but also enjoyable, attractive and appealing. Participants in C_A reported they would be more likely to come back in the next six months for a similar visit than those in C_N (6.4 VS 4.0, p-value≪0.01) or to recommend the exhibition to others (6.87 VS 4.0, p-value≪0.01). Furthermore, participants reported they found the application visually pleasing and easy to use.

[1] AI-based image generator: https://openai.com/dall-e-2

Memorization performance. Results displayed benefits of the app when the users were asked on details of artworks related to riddles (48% VS 17%, p-value<0.01). We found no differences when details were not related to riddles (41% VS 33%, p-value=0.19). Concerning description, we observed a hindrance tendency when using the app (62% VS 79%, p-value=0.052).

Conclusion & Future works

We proposed and evaluated Star Catcher, a mobile app aiming to improve user engagement for art exhibitions. A user study highlighted that the application was really appreciated by our participants. They reported a high intention to come back and recommend the tour to others, showing a great interest of the app for museums willing to improve visitors number compared to participants in control group. Most participants also told us they would have like to have more information on details of the artworks related to riddles. This is encouraging feedback since it suggests that our participants were truly engaged with the art exhibition.

Our results also showed the application enhances memorization of artworks, but lessen memorization of descriptions. We suggest two possible explanations for this result. A first explanation could be that the application focused the attention of participants more on the artworks than their description. A second explanation could be that our participants were not used to art exhibitions, as they reported on average usually going to museums less than once a semester. They would then naturally focus less on learning about artworks than real visitors. Our next step is to deploy Star Catcher in a real art exhibition, with real visitors. This will allow us to study how the application influences memorization when people actually come for the art exhibition, using riddles related to meaningful details of exposed artworks.

References

McLean, G., & Wilson, A. 2019. Shopping in the digital world: Examining customer engagement through augmented reality mobile applications. Computers in Human Behavior, 101. pp. 210–224. https://doi.org/10.1016/j.chb.2019.07.002

Gong, Z., Wang, R., & Xia, G. 2022. Augmented Reality (AR) as a Tool for Engaging Museum Experience: A Case Study on Chinese Art Pieces. Digital, 2(1). pp. 33–45. https://doi.org/10.3390/digital2010002

Nguyen, D., & Meixner, G. 2019. Gamified augmented reality training for an assembly task: A study about user engagement. In: Fed. Conf. on Computer Science and Information Systems, Leipzig, Germany. pp. 901–904. https://ieeexplore.ieee.org/abstract/document/8860007

Laugwitz, B., Held, T., & Schrepp, M. 2008. Construction and Evaluation of a User Experience Questionnaire. In: HCI and Usability for Education and Work, Graz, Austria. pp. 63–76. https://doi.org/10.1007/978-3-540-89350-9_6

Retrospective on the Development and Application of the Dialog Simulation Format in VR

Lidia Yatluk

[1] *University of Groningen, The Netherlands*

Corresponding author: LidiaOutlook@gmail.com

Keywords: Virtual Reality, Soft Skills, Workforce Training, Learning Format

Abstract

The text is devoted to the retrospective of the development and application of the dialog simulation format in Modum Lab. It describes the initial ideas, user experience, and application areas at different stages of the product, changing priorities, and the results the format has achieved. The main feature of the format's transformation is the shift from a realism orientation to standardization and compatibility with traditional e-learning. The paper concludes with a comparison with the similar developments of other startups and perspectives of the format.

Introduction

The concept of practicing communication skills in virtual reality (VR) has existed since early consumer devices. The most common applications include foreign language training (Mondly, VARVARA, Immerse), public speaking (Virtual Speech, BeFearless, Ovation), and medical professional communication training (e.g., Lok et al 2006; overview - Alivi et al 2023).

Workplace training using VR gained traction later. In 2018, Oculus Connect 5 introduced a Walmart training program, featuring Black Friday stress training. In 2020, PricewaterhouseCoopers released a significant report on workforce training in collaboration with Talespin (2020) for the corporate sector.

Dialogue simulation format in the observed company was shaping since 2017 until 2021 with recurrent user experience research and business effectiveness evaluation. This article aims to describe the process of integrating dialog simulations and share best practices based on production of 12 complex projects across banking, retail, IT, and process industries.

Dialogue simulation format

A dialog simulation is a VR training format where voice communication occurs between bots and a user, practicing communication skills using a scenario tree. Think of it like a gym, not perfectly replicating real situations but focusing on specific aspects for practice, much like a simulator. Similar to fitness equipment, they require repeated use for skill development.

Technically, a dialog simulation is a two-level system involving speech recognition and classifying statements based on rules. Bot responses are pre-scripted and recorded. As this format predates the rise of large language models, rules are manually defined by script developers. Conditions could involve word sets, their positions, and differences from reference words (allowing for changes in endings and suffixes).

Dialog simulations target employees aiming to enhance specific communication skills while learning corporate standards.

History of changes

The desire for realism

In 2017, the key motive driving dialog simulation development was realism. This involved creating diverse animations for bots and crafting scenarios with a wide array of user responses for an authentic dialogue experience.

Crafting dialog simulations with extensive user discretion proved time-intensive. Development demanded detailed input from trainers familiar with teaching the skill and multiple focus-groups. Ideal, acceptable answers, and common mistakes needed definition prior to course creation.

User experience yielded contradictions. While freedom facilitated quick immersion – with skeptical users engaging within 3 minutes – it sometimes led to unvarying bot reactions, causing frustration due to repeated misunderstanding.

A significant challenge in realism orientation was integrating dialog simulation with other skill practice. One project necessitated inspecting a car and conversing with customers concurrently – clarifying requests, learning breakdown history, and boosting sales. This complex scenario was further complicated by potential user misinterpretations in the virtual enviroment.

The desire for predictability

In 2019, predictability gained importance in dialog simulation development, aiming for more accurate development time estimates and consistent bot interactions. This shift led to reduced user freedom, favoring deterministic scenarios over freely conversational ones.

Three main types emerged: dialog exercises, linear dialogs, and response-option dialogs. Dialog exercises reinforce specific conversational techniques, while linear dialogs provide step-by-step guidance. Response-option dialogs teach skills like feedback and meetings, using recommended topics rather than fixed responses.

These new dialog types altered user experience. Initial enthusiasm waned, and immersion slightly diminished (measured through surveys and interviews). However, accessibility improved, and alignment with learning analytics based on simulation results increased. Overall usage frequency in companies also rose.

Development predictability improved. Deterministic scenarios reduced MVP development time, minimizing focus groups and refinement iterations.

The transition to omnichannel experiences

In 2020, even before the pandemic, it became evident that immersive VR was not suitable for all user scenarios. This led to the idea of blending small dialog simulations on smartphones and browsers with longer immersive experiences in VR.

Accessibility took center stage during this period. Dialog simulations were designed to be lightweight, running smoothly on different devices. Interfaces were optimized for mobile use, removing diegetic elements for consistency across devices. Three-dimensional training rooms were replaced with simpler backdrops. Dialog simulations in this phase resembled earlier computer-based versions.

With the loss of immersion, training duration was shortened to maintain engagement. Sessions were divided into brief simulations for focused practice.

User experience now varied based on the simulation's usage. VR could still be enjoyable in certain contexts, while on other devices, the experience felt routine.

Return to e-learning

In 2022, dialog simulations continued their shift towards classic e-learning. The focal point was connectivity encompassing learning experiences, analytics, and technical integration. Dialog simulations started being packaged in the standard SCORM format and seamlessly integrated into popular enterprise learning platforms.

Simplifying adaptation to other formats required an intuitive editing feature, leading to the active use of an in-house scenario builder Verba. This tool allowed basic scenarios to be visually assembled using a small set of programming commands. More open simulations and similar ones were phased out due to complexities in this new format.

The primary dialog simulation type became concise and task-specific, focusing on highly refined algorithms rather than global soft skill development. Its purpose was to guide users through distinct situations with precision. This approach found prominence in retail for training in customer consultation, upselling, and customer service.

Consequently, the user experience resembled that of dialog exercisers or computer simulations. Training garnered average ratings, with higher scores when storytelling was well-crafted and seamlessly integrated into the overall training context.

Conclusion

From 2017 to 2023, dialog simulations evolved from immersive projects with animations to microlearning conversational skills in training programs. This shift coincided with market frustration in VR and metaverse for various purposes.

Dialog simulations have become business efficient for the developer and the customer. They are developed in a predictable amount of time, require relatively standard pedagogical design qualifications, and work with a high-demand area for training. For the customer, dialog simulations have become a comprehensible, integrable format that can verifiably improve the learning of corporate standards.

In terms of utilizing the capabilities and specificity of VR, it might be said that they have been abandoned. The original reliance was on isolation from external stimuli (everything is done in a VR headset), emotional immersion (due to the freedom of conversation and the features of bots), a sense of presence (due to the elaborate three-dimensional environment), and a sense of authenticity of experience (due to the connectivity of all elements). Dialog simulations have evolved into regular computer simulations that are sometimes available in VR. Previously planned changes to bot animations, diegetic cueing interfaces, and pose-shifting simulations were canceled.

Comparing these developments with publicly available information about similar startups like Talespin and Bodyswap reveals notable differences in their trajectories. Talespin has largely followed a path of simple algorithmic scripts, hints, and dividing training into smaller segments. They have also developed their own simple script editor known as "CoPilot Designer." However, they have retained a strong focus on detailed conversational bots, presenting virtual humans as a core component of their platform on their official website. Their primary aim is to cultivate flexible skills rather than distinct standards. Bodyswap, on the other hand, has also employed a strategy of fixed options within dialogues but has enriched the experience by incorporating reflection in the form of listening to one's own responses from the perspective of the interlocutor, as evidenced in their "Bodyswap Learning Framework."

All three paths share a common trend of simplification and predictability, moving away from simulating full-fledged human communication as was the initial goal during the early stages of adoption. This transition aligns with the typical phase of technological diffusion adaptation (Dearing and Cox 2018).

Complex ideas from 2017 are now achievable through language models (ChatGPT, LLaMA), voice generation tools (Play.ht, Murf.AI), and 3D object generation (Spline AI, Masterpiece Studio), enabling possibility of original VR fidelity within the business sustainability frame.

References

Aliwi, I., Schot, V., Carrabba, M., Duong, P., Shievano, S., Caputo, M., Wray, J., de Vecchi, A., & Biglino, G. 2023. The Role of Immersive Virtual Reality and Augmented Reality in Medical Communication: A Scoping Review. Journal of patient experience 10. https://doi.org/10.1177/23743735231171562

Bodyswaps: VR Soft Skills Training Provider. https://bodyswaps.co/, accessed 18.8.2023.

Bodyswap: Learning framework. https://bodyswaps.co/wp-content/uploads/2021/03/Bodyswaps-Learning-Framework.pdf, accessed 6.10.2023

Dearing, J.W., Cox, J.G. 2018. Diffusion of Innovations Theory, Principles, And Practice. Health Affairs. 37(2):183–190. https://doi.org/:10.1377/hlthaff.2017.1104.

Dialogue simulation. About the format. 2019. https://www.youtube.com/watch?v=jydpf2OvBhU, accessed 17.8.2023

Lok, B, Ferdig, R.E., Raij, A, Johnsen, K., Dickerson, R., Coutts, J., Stevens, A., Lind, D.S. 2006. Applying virtual reality in medical communication education: current findings and potential teaching and learning benefits of immersive virtual patients. Virtual Reality 10(3),185–195. https://doi.org/10.1007/s10055-006-0037-3.

Modum Lab unveils e-commerce development product. TAdviser.ru. https://tadviser.com/index.php/Product:Modum_Verba_Dialog_Simulation_System, accessed 17.8.2023

Oculus Connect 5 | Training, Transformed: Walmart Explores VR at Scale. 2018. https://www.youtube.com/watch?v=xTvTv5CU8XE, accessed 17.8.2023

PricewaterhouseCoopers. PwC's study into the effectiveness of VR for soft skills training. PwC. []. https://www.pwc.co.uk/issues/emerging-technologies/metaverse-technologies/study-into-vr-training-effectiveness.html. accessed 17.8.2023

Talespin - Enterprise VR, AR & AI. https://www.talespin.com/, accessed 18.8.2023.

"*Zauberbuch*" – An Interactive, Projective AR-Based Approach toShow Content in a Blank Book

Nick Weidensager, Sven Winkler, Christian Fuchs, and Franziska Klimant

[1] *Chemnitz University of Technology, Institute for Machine Tools and Production Processes, Chemnitz, Germany*
Corresponding author: nick.weidensager@mb.tu-chemnitz.de

Keywords: Projective Augmented Reality, OCR, Tesseract, Educational Application, Book

Abstract

Projective Augmented Reality sees use in a variety of application areas to support users at certain tasks or to create better ways of visualization, especially where wearing special equipment would be unpractical or the number of users varies greatly. In this paper, the development of an interactive, projective AR-based book-prototype, which is being used in a public university library, and so far gathered user feedback, is described.

Introduction

Applications of projective Augmented Reality (AR) cover a big spectrum of daily life. Use cases that are researched so far range from industrial topics like design and construction of buildings (Xiang et al., 2021) or shared human-robot workplaces (Vogel et al., 2020) over medical data analysis (Heinrich et al., 2020) to educational (Radu, 2012) and artistic applications (Pletcher et al., 2020). In most applications, projective AR surpasses handheld or HMD-based AR, because the user does not need to wear or hold certain equipment and multiple people can view the AR content at the same time.

Within the project "Zauberbuch" (translated: magic book) lead by the university library of the Chemnitz University of Technology, a prototype was developed that couples projective AR and a real but mostly empty book to introduce an interactive way of presenting information for visitors of the university library. Under the name "*Zauberbuch*", visitors can use the prototype since June 2023. The name "Zauberbuch" was chosen to avoid mix up with the magic book developed by Billinghurst et al. (2000), which features a normal book with HMD-based and handheld AR. Contrary to their approach, projective AR is used because a large number of users with unknown limitations to wearing equipment is expected to use the "Zauberbuch". The following section will describe the development of the prototype system and software, including user feedback, gathered so far.

Development of the prototype

Since the "*Zauberbuch*" is presented in a public library and should be accessible for all users, the idea was to use a real book with only page numbers printed on the pages. That way, users can intuitively change shown information by flipping the pages of the book. Compared to using gesture-based interaction techniques, there is less room for operation errors, for example by moving the hand too slow or too fast. Gesture-based interaction also has the problem of users often not knowing where to move their hands, so that a certain action triggers unintentionally. For recognition of the currently opened page via page number, the open access optical character recognition (OCR) software "Tesseract" was used and modified to fit the usage in a video stream. Hence, the "*Zauberbuch*" consists of a beamer, a webcam, a PC and the book itself. A custom fit table to hold and secure all components completes the developed prototype. The prototype and its components are shown in Figure 2.

The book as heart of the "Zauberbuch"

As described before, the projection space of the prototype is a blank book that only shows printed page numbers at the top of each double page. A special bookbinding was used to get the surface as flat as possible. Because of the special binding and the high paper thickness, the number of double pages is limited to 24, which means a limited amount of virtual content. However, since the "*Zauberbuch*" is intended as a part of a themed exhibition and 24 pages can hold a lot of information, this limitation is insignificant. The book itself is screwed on the table to prevent unwanted displacement and thus eliminating the need for recalibration of the projector.

Figure 2. Left: Projective AR prototype "Zauberbuch", Top Right: The closed book from the users point of view, Bottom Right: Content of the first double page displayed on the book.

Software development

The software was self-developed within the game engine "Unity"[2] and runs on a mini-PC that is located inside the table. In order to determine the currently opened page the "Unity" package "Tesseract", which processes a single image and returns recognized text, was implemented. This single image recognition is not focused on time performance and would lead to issues with the simultaneous presentation of content, which is why the base package had to be adapted. The development of the software was mainly influenced by the following requirements: the positioning of the camera and the performance of the application in general. The used camera should be out of view of the current user. Additionally should its tilt, when looking at the book, be small, to minimize errors in text recognition. Therefore, the best position would be behind the beamer, facing in the same direction. This leads to the camera capturing a lot of space beside the page number, especially the content that is projected on the book. By trimming every video-frame that is fed into "Tesseract", those interfering areas can be removed. This also assists in meeting the second requirement, because the main source of computing time is "Tesseract's" image processing, which is reduced significantly by reducing the resolution of the input image. Additionally not every recognized page number should lead to a change of shown content, because sometimes numbers or even background noise are falsely recognized. Thus, a mechanism was integrated to only change content for page numbers that were recognized multiple times in a row, increasing computing time again.

Frame rate of the "Zauberbuch"

Low computing time is necessary to achieve an acceptable performance, in this case meaning a high enough and stable frame rate with which the AR content is shown. The study of Mackin et al. (2015) took a look at subjective video quality at different frame rates. Their study shows a mean opinion score of around 70 % for 60 frames per second (FPS) videos and only 55 % for 30 FPS videos. That means perceived quality increased by roughly 15 % by having twice the frame rate, whereas an increase from 60 to 120 FPS only meant 5 % increase in mean opinion score. Knowing that, the target frame rate for the "*Zauberbuch*" should be near or above 60 FPS, so users will not perceive the content as stuttering. By trimming the size of the input images, the computing time of the "Tesseract" package was reduced from between 100 to 250 ms, which would equate to 10 to 4 FPS drops, to less than 40 ms (> 25 FPS), thus preventing visible drops in frame rate when running the "Tesseract" algorithm multiple times per second.

Storing and organizing the virtual content

The content of the "*Zauberbuch*" is stored in a database and can be shown in two ways. In the first layout, the double page is vertically split into two pages with the left side displaying text and image subtext and the right site featuring a portrait image. The second layout supports landscape images covering most of the top half of the double page, excluding side margins, with text shown beneath the image. All data is loaded on startup to circumvent loading time on runtime.

In the final setup of the prototype, 10 consecutive recognitions of a page number are necessary to trigger the content change. Setting the rate of processed images to five per second ensures a stable performance of around 60 FPS. Hence, roughly two to three seconds after opening a page the corresponding content will be shown.

[2] www.unity.com

User Feedback

Since the inauguration, the "Zauberbuch" was accessible for all visitors of the library. Although no specific surveys were being made, the overall feedback was very positive regarding the robustness of the prototype. Some visitors commented on the long duration of page recognition, which could be reduced after further testing. Another critique point was missing instructional information, because the first page is empty and thus no content will be shown. In the near future, a methodological survey will be provided at the prototype to better analyse user acceptance of the presented interaction system on a voluntary basis.

Conclusion

In this paper, a prototype for an interactive, projective AR-based book and its development have been described. By using the open access OCR software "Tesseract", page numbers on otherwise empty pages are recognized and corresponding content, stored in a database, is displayed by a projector. Thus allowing users to change shown content by flipping the pages of the book. After focusing on a robust number recognition, user feedback was mainly positive, but some found the recognition time to be too long. Improvement of recognition time alongside new features and instructions will be worked on in the near future.

Acknowledgements

This project has received funding from University of Technology Chemnitz within the TUCculture2025-funding programme.

References

Billinghurst, M., Cambell, S., Poupyrev, I., Kato, K. T. H., Chinthammit, W., & Hendrickson, D. 2000. Magic book: Exploring transitions incollaborative arinterfaces. In Proc. of SIGGRAPH 87.

Heinrich, F., Bornemann, K., Lawonn, K., & Hansen, C. 2020. Interacting with medical volume data in projective augmented reality. In Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd Int. Conf. Part III, pp. 429-439. Lima, Peru, October 4–8, Springer International Publishing.

Mackin, A., Zhang, F., & Bull, D. R. 2015. A study of subjective video quality at various frame rates. In 2015 IEEE International Conference on Image Processing (ICIP), pp. 3407-3411, https://doi.org/10.1109/ICIP.2015.7351436.

Plecher, D. A., Ulschmid, A., Kaiser, T., & Klinker, G. 2020. Projective augmented reality in a museum: development and evaluation of an interactive application. In: ICAT-EGVE 2020 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments. https://doi.org/10.2312/egve.20201258

Radu, I. 2012. Why should my students use AR? A comparative review of the educational impacts of augmented-reality. IEEE Int. Symp. Mixed Augm. Reality (ISMAR), pp. 313-314. https://doi.org/10.1109/ISMAR.2012.6402590

Vogel, C., Schulenburg, E., & Elkmann, N. 2020. Projective-AR assistance system for shared human-robot workplaces in industrial applications. In 2020 25th IEEE Int. Conference Emerg. Technol. Factory Autom. (ETFA), 1, 1259-1262. https://doi.org/10.1109/ETFA46521.2020.9211953

Xiang, S., Wang, R., & Feng, C. 2021. Mobile projective augmented reality for collaborative robots in construction. Automation in Construction, 127, 103704. https://doi.org/10.1016/j.autcon.2021.103704

BIM Digital Twin Visualization in MR – Pilot Evaluation

Kaj Helin, Timo Kuula, Vladimir Goriachev, and Jaakko Karjalainen

VTT Technical Research Centre of Finland Ltd, Tampere/Espoo, Finland

Corresponding author: kaj.helin@vtt.fi

Keywords: BIM, Digital Twin, Mixed Reality, Pilot Evaluation

Introduction

This extended abstract introduces describes Mixed Reality (MR) tool called BIM@Construction, which is part of GUI concepts and visualizes digital twin data on a construction site. The GUI concept has been introduced earlier (Helin et al., 2022). The pilot tests for BIM@Construction were conducted in HRS construction site in Lausanne, Switzerland in November 24, 2022. Pilot tests result are presented in this paper. Work has been done in a European Commission funded H2020 project called "BIMprove - Improving Building Information Modelling by Realtime Tracing of Construction Processes"[3].

BIM@Construction – MR tool

Main idea behind BIM@Construction concept application is to use BIM as digital twin data in immersive XR devices at the construction site. While MR devices like Microsoft HoloLens 2[4] are already built for interaction with 3D models (as shown in Figure 3), scale of those models is usually dictated by the FOV (Field of View) of the headset. In order to overlay digital twin information on top of the real building, additional pivot objects should be created. Naturally, pivot object placement should match the placement of the fiducial markers in the real world. That way the digital twin and its real-world counterpart will be located in the same coordinate system. BIM@Construction includes several tools, which user can use in 1st person mode, like notes, measure tools, warning signs etc.

[3] BIMprove, CORDIS, https://cordis.europa.eu/project/id/958450, last accessed 2023/06/28.
[4] Microsoft HoloLens mixed reality platform, https://www.microsoft.com/en-us/HoloLens, accessed 28.6.2023

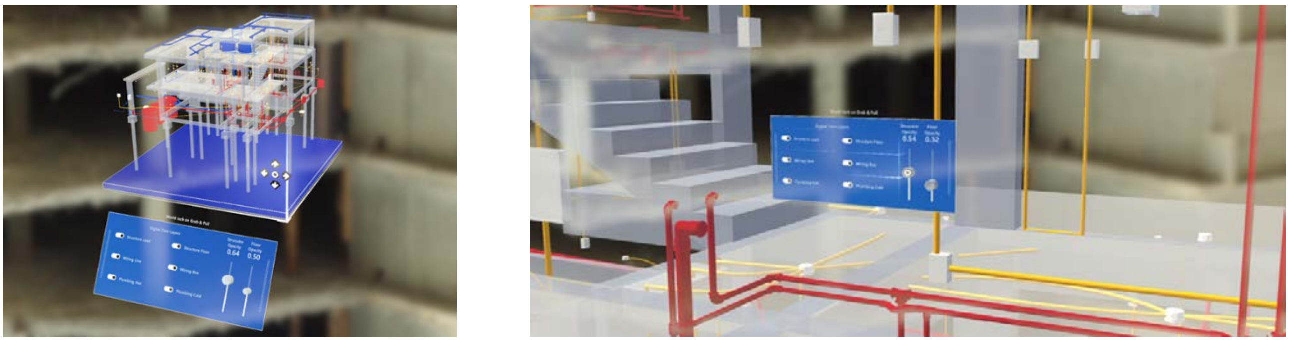


Figure 3. Immersive AR version of BIM@Construction. Left: 3rd person view. Right: 1st person view.

Pilot Tests

The pilot tests for BIM@Construction were conducted in HRS construction site in Lausanne, Switzerland in November 24, 2022. More specifically, the test space was a basement room of an apartment house under construction. The test system and space were set up during the previous day. The goal of the tests was to collect feedback from actual construction site personnel in the final phase of the development.

Altogether five (5) HRS employees participated in the tests, all of them representing the supervisor or manager level. The test persons performed the test individually. The tests were prepared, instructed and guided by researchers from VTT (see Figure 4).

Figure 4. Example of pilot tests. Person is testing BIM@Construction and HF expert is guiding the test and observing.

Use case

BIM@Construction can be viewed in first-person mode, that is, in the immersive mode (see Figure 5), in which the viewer is as if inside the building. User can do various tasks inside the building, both by deciding what details to see and also by adding more information to the building with digital means. It can also be perceived in the third-person mode, as if outside the building, so that the user can rotate the building model, look at it closer or further as needed.

Figure 5. Immersive mode. Left: Checking piping. Right: Using measurement tool.

During the actual test performance, the users wore HoloLens glasses and performed tasks in the basement room following the instructions from the test supervisor. The task included all main features of the BIM@Construction MR-system.

Evaluation methods

The test participants filled in System Usability Scale (SUS) questionnaire. The SUS scores calculated from individual questionnaires represent the system usability. SUS yields a single number representing a composite measure of the overall usability of the system being studied. Scores for individual questionnaire items are not meaningful on their own (Brooke, 1996). SUS scores have a range of 0 to 100. The users filled in the SUS questionnaire immediately after the test performance.

Results

According to validation studies (Bangor et al., 2009; Brooke, 2013), the SUS score starting from 68-70 represents the level of acceptable system usability. Furthermore, the suggested acceptability ranges are: 0-50 not acceptable; 50-70 marginal; 70- acceptable.

SUS scores:

Test participant	SUS score
1	90
2	73
3	83
4	43
5	83

SUS score average: 74

Based on the SUS score average (74), the system usability was evaluated as acceptable. All expect one of the individual test participants evaluated the usability as acceptable, and three of the users gave a score over 80, which indicates very good usability. It should be noted that the result does not provide statistical significance due to low number of test participants.

As one user gave a score less than 50, meaning that usability is not on acceptable level, this should be further elaborated. The score was not at all in line with the rest of the scores. It should be noted that the particular user suffered from simulator sickness symptoms during and after the test.

Conclusion

Based on the pilot test results, developed BIM@Construction MR tool concept seems to be working properly, and its usability has reached acceptable level. It also allows user to see various information levels on-site based on user roles.

Acknowledgment

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 958450.

References

Bangor, A., Kortum, P. & Miller, J. 2009. Determining What Individual SUS Scores mean: Adding an Adjective Rating Scale. Journal of Usability Studies. 4(3, 114–123.

Brooke, J. 1996. "SUS: a 'quick and dirty' usability scale. In Usability evaluation in industry. P. W. Jordan, B. Thomas, B.A. Weerdmeester and I. L. McLelland (Eds.) Taylor and Francis Ltd. London, UK.

Brooke, J. 2013. SUS: A Retrospective. Journal of Usability Studies 8(2), 29–40.

Helin, K., Goriachev, V., Karjalainen, J., Kuula, T., Liinasuo, M., & Aust, M. 2022. BIM digital twin visualization in XR – Novel GUI concept. Proc. EuroXR Conference, pp. 59-62. VTT Technical Research Centre of Finland. https://doi.org/10.32040/2242-122X.2022.T408

Appendix

Video from the user tests: https://youtu.be/nveo16orX1w

The Kitchen of the Future: Testing User Experience Through Virtual Reality

María Plaza[1], Nerea Doncel[1], Ana Serrano[1], Sergio Llorente[2], Julio Rivera[2], Francisco Javier Ester[2], and Diego Gutiérrez[1]

[1]*Graphics and Imaging Lab - University of Zaragoza I3A, Spain*
[2]*Department of Research and Development, Bosch-Siemens Home Appliances Group, Spain*

Corresponding author: mplaza@unizar.es

Keywords: Virtual Reality (VR), User Experience (UX) Test, The Kitchen of the Future

The evolution of virtual reality (VR) has led to a range of innovative applications in various fields, and the culinary household appliance industry is no exception. In our pursuit of exploring new possibilities for the kitchen of the future and within the context of a collaborative research project with a leading company in the sector like BSH, we have posed the question: How can we harness virtual reality (VR) to enhance user testing processes?

This contribution aims to explore the potential of VR as an innovative tool to enhance evaluation processes in the development of the kitchen of the future. We believe that our research can contribute to understanding how VR can influence user interactions with products and ultimately drive innovation and continuous improvement in the field of user experience (UX) testing for the kitchen of the future.

Introduction

User experience (UX) testing is paramount in the development of new products and services. Traditionally, market research studies and user testing have relied on conventional methods such as questionnaires, face-to-face interviews and direct observation. However, these approaches may have limitations in terms of reaching a wide range of users, conducting test across different countries, precisely understanding user attention during interactions with products or the substantial expense and effort involved in constructing a physical prototype, which may be completely rebuilt in the event of any necessary modifications.

VR provides a unique opportunity to overcome these limitations by offering an immersive, portable and, to a greater extent, automatable environment. Although there are currently challenges regarding interaction in virtual environments, we believe that VR can bring valuable improvements to evaluation processes.

In this contribution, we present our approach to utilizing VR in UX testing within the context of the kitchen of the future. We explore the advantages and limitations of this technology in evaluating UX and describe the experimental design and preliminary results of our study. Additionally, we discuss the practical implications and potential future research directions in this emerging field.

Methodology

A study was conducted based on a test previously conducted by the company BSH, which involved a medium level of interaction between the user and the product. The test involved evaluating five types of controls for a new induction cooktop solution. In this test, there was physical/tactile interaction with the prototypes, although they were not fully functional since the goal was not to test the functionality of each prototype but to understand user preference. This aspect is crucial as we needed to explore how to address or transfer this interaction to VR in this initial approach, without the physical prototypes of the controls.

The primary objective of the study is to ascertain the feasibility of conducting the test in VR. For this purpose, an initial pilot test was carried out and once the viability of the test was confirmed, we design the final test.

In this initial pilot of the test, the exact same test conducted by BSH was replicated in VR, modeling the same space as the original test and reproducing the prototypes with their original manufacturing flaws.

Figure 6. Left: representation of the test conducted by BSH. Right: virtual reality modelling of the same test.

During the test design, the importance of combining real physical supports during the use of VR headsets was recognized, allowing users to rest on a tangible surface that matched the one seen through the goggles, promoting a secure interaction. Additionally, it was necessary to acquaint participants, especially those unfamiliar with VR, with wearing and using VR devices. In order to ensure consistent results, the users were first invited to freely explore a different scenario from the test and once they were familiarized, then they saw the test environment.

In the final test, the advantages offered by virtual reality were leveraged, such as creating a realistic kitchen immersive environment, improvements and updates to the prototypes (e.g., interfaces), which could not be carried out in the original test due to manufacturing costs and time constraints.

The original test methodology was divided into four blocks, which were carried out in the same way in both VR and the original test, involving constant dialogue between the user and the interviewer.

The first block focused on the user's lifestyle in relation to their cooking habits. The second block addressed how the user perceives the verbally described product and its ideal control. The third block involved the interaction phase, and the last block consisted of an evaluation. It was in this third block that the user put on the VR glasses and carried out the UX test until its completion. Finally, a fifth block was added in which, instead of presenting the controls one by one, all of them were shown simultaneously to encourage creative brainstorming and ensure that no potential solution was overlooked.

Figure 7. Immersive environment of the final test.

Figure 8. Left: diagram of the dynamic of the test. Right: real picture of the test.

Results and conclusion

The qualitative results obtained from the tests indicate that the identified pains and gains for each concept follow a similar trend, revealing that the concepts are understood in the same way in both reality and VR. This confirms the viability of conducting UX tests in VR. The first quantitative results, reported with a sample size of ten users, indicate that preferences remain consistent, with the same favorite and least favorite concepts, and the remaining concepts maintain the same order of preference with a single exception. We hypothesize that this could be attributed to various factors, such as differences in user characteristics, among others, which could be explored in future research.

Regarding the overall findings, we observe the brain's ease of adaptation. Users had a genuine sense of immersion in the environment shown to them. Furthermore, having a physical support helps enhance the immersive experience and user safety. The physical support worked even if its height was not an exact match to what was seen through the goggles. Initially, users noticed the difference, but they adapted within seconds once they touched it. Along the same line, the training to become familiar with the devices took less than a minute in this case. On the other hand, people did not experience dizziness or fatigue, a critical consideration given the potential challenges in VR experiments. Lastly, in some cases, hearing the interviewer without seeing them led to confusion. This aspect is open for further exploration in future studies.

In conclusion, both the Lab and BSH agreed on the viability of conducting VR tests and its potential advantages, such as the ease and cost-effectiveness of modifying virtual prototypes. Moreover, it provides the opportunity to conduct tests in various locations and reach a broader audience more easily. Since virtual reality is still in the process of improvement and development, and has not yet reached the point where it can be fully automated to the extent that anyone with VR goggles can perform a test without the assistance of an interviewer, our future work will focus on exploring ways to automate VR testing, determining the optimal levels of interaction for its use, and implementing interaction enhancements, such as hand tracking.

Acknowledgements

We would like to thank all the people who have been involved in the design of the test and the test itself. We would also like to thank BSH for its commitment to innovation and research with the collaborative project that gives rise to this research.

This work is part of project CPP2021-008938, funded by MCIN/AEI/10.13039/501100011033 con-funded by European Union's NextGenerationEU/PRTR and in part by BSH Home Appliances Spain. And by the Government of Aragon's Departamento de Ciencia, Universidad y Sociedad del Conocimiento through the Reference Research Group "Graphics and Imaging Lab (ref T34_23R).

References

Bruno, F., & Muzzupappa, M. 2010. Product interface design: A participatory approach based on virtual reality. International journal of human-computer studies, 68(5), 254-269. https://doi.org/10.1016/j.ijhcs.2009.12.004

de Freitas, F. V., Gomes, M. V. M., & Winkler, I. 2022. Benefits and challenges of virtual-reality-based industrial usability testing and design reviews: A patents landscape and literature review. Applied Sciences, 12(3), 1755. https://doi.org/10.3390/app12031755

Falcão, C. S., & Soares, M. M. 2013. Application of virtual reality technologies in consumer product usability. In Design, User Experience, and Usability. Web, Mobile, and Product Design: Proc 2nd Int. Conf. DUXU 2013, Part IV 2. 352-351.Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21-26, Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-39253-5_37

Rebelo, F., Noriega, P., Duarte, E., & Soares, M. 2012. Using virtual reality to assess user experience. Human factors, 54(6), 964-982. https://doi.org/10.1177/001872081246500

Enhancing Geometric Coherence in Digital Twins through Localized Augmented Reality-Based Positioning

Abdelhadi Lammini, Frederic Noel, Romain Pinquie, and Gilles Foucault

Grenoble INP, France

Corresponding author: abdelhadi.lammini@grenoble-inp.fr

Keywords: Geometric Coherence, Digital Twins, Augmented Reality, Positioning, Quality Test Machines

Introduction

Geometric coherence plays a pivotal role in consistently capturing and representing the position and orientation of digital twins' snapshots in the context of flexible systems such as assemblies of parts and furniture in a factory room. While efficient indoor positioning systems exist, deploying them at a factory scale to manage every piece of equipment is challenging. Our study seeks to introduce an alternative method. We propose a novel approach based on human detection of geometry changes with augmented reality support, aiming to enhance the coherence of coordinate systems within snapshots of the digital twin's lifecycle. Our particular focus lies in leveraging localized augmented reality (AR)-based positioning techniques to address the challenges encountered by existing methods. These challenges revolve around consistently positioning the reference coordinate system of geometric snapshots taken at distant dates within the same building, ultimately advancing the accuracy of hologram alignment on HoloLens2 glasses.

Methodology: Localized Augmented Reality-Based Positioning

Hologram positioning relies on two methods: image targets like QR codes and area targets like Azure Spatial Anchor (ASA). Image targets can't correct Inertial Measurement Unit (IMU) sensor drift, causing positioning inaccuracies over time. In contrast, area targets use AR devices like HoloLens 2 to scan the environment for precise placement, correcting sensor errors by overlaying reference scans with spatial mapping data using algorithms like Iterative Closest Point (ICP) or Random Sample Consensus (RANSAC) (Bo et al.).

Area target demand significant computational resources and relays on the network. The Triplane Anchoring Method (TAM) offers an innovative approach for dynamic element positioning using HoloLens device-based positioning. TAM establishes a persistent coordinate system using non-parallel room surfaces, ensuring stability in various indoor scenarios. It enables efficient tracking of dynamic parts without cloud computing, enhancing geometric coherence over time. TAM selects surfaces for anchoring, ensuring consistency across snapshots and seamless transitions between positioning systems. This process enhances persistent object positioning for dynamic digital twin components and static room surfaces on HoloLens 2. The steps of TAM execution are as follow:

1. Surface Identification: Vocally selecting nine large distant triangles (three per surface) from the visible spatial mapping of non-parallel surfaces in the scene to minimize orientation errors.
2. Mean Position, Orientation, and Normal Calculation: Once three triangles are chosen, a new big triangle is created from each of their centroid, defining the orientation of the surface.
3. Surface Intersection and Coordinate System Creation: Computing intersection points among selected surfaces to determine the center and axes of the coordinate system. The center point serves as the origin (0,0,0), and the normals define the X, Y, and Z axes of the coordinate system.
4. Object Placement and Saving: The user positions the desired object within the established coordinate system, and the object's position is recorded and saved on the HoloLens device.
5. Coordinate System Storage: Saving the created coordinate system, including its center point and axes, on the HoloLens device for future reference.
6. Uploading Object to Scene: Uploading the saved object with its associated coordinate system into the augmented reality scene for precise placement.
7. New Coordinate System Selection and Alignment: Repeating steps 1 to 3 for the same selected surfaces to create a new coordinate system, followed by aligning the uploaded coordinate system with the new one for object placement.

During Hololens2's spatial mapping, selected surfaces for new coordinate systems automatically update to counter IMU sensor drift, ensuring stable object positioning. The algorithm identifies congruent normal triangles to calculate these updated coordinates.

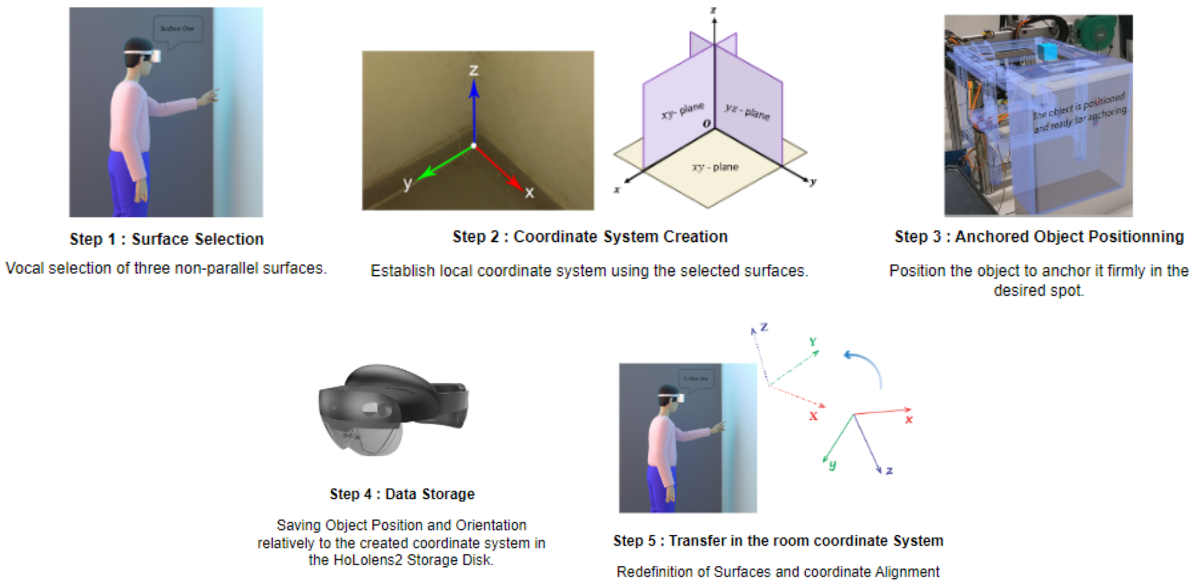


Figure 9. *Localization Process for Geometric Coherence Enhancement using TAM method.*

Use Case: Enhancing Accuracy in Quality Test Machines

To illustrate the enhancement of precision through AR-based positioning, the TAM, and ASA methods are combined with an image target solution, namely a QR code. Each part is outfitted with a QR code, serving as a reliable source of positional information upon scanning. This information is pivotal for comparing the performance of the TAM and ASA methods.

To simplify the detection of deviations for the observer's convenience, we opt to represent the sensor's 3D model using a cube as shown in Figure 10.

We assume that the initial scanning of the QR code has a low positioning error. Any discrepancies in the distances between the QR object and the TAM or ASA objects are neglected, providing the deviation in the x, y, and z axes.

In this comparison, we focus on anchor positions determined by Azure's method and the TAM. Precision assessment between TAM and ASA methods involves several steps: (a) Identifying the anchor position using both methods, (b) scanning the QR code, and (c) instantly comparing distances between the QR code object and previously anchored positions obtained through ASA and TAM to minimize deviations.

Figure 9 illustrates the sequence of anchoring procedures involving both methods, followed by anchor identification and the subsequent deviation detection process.

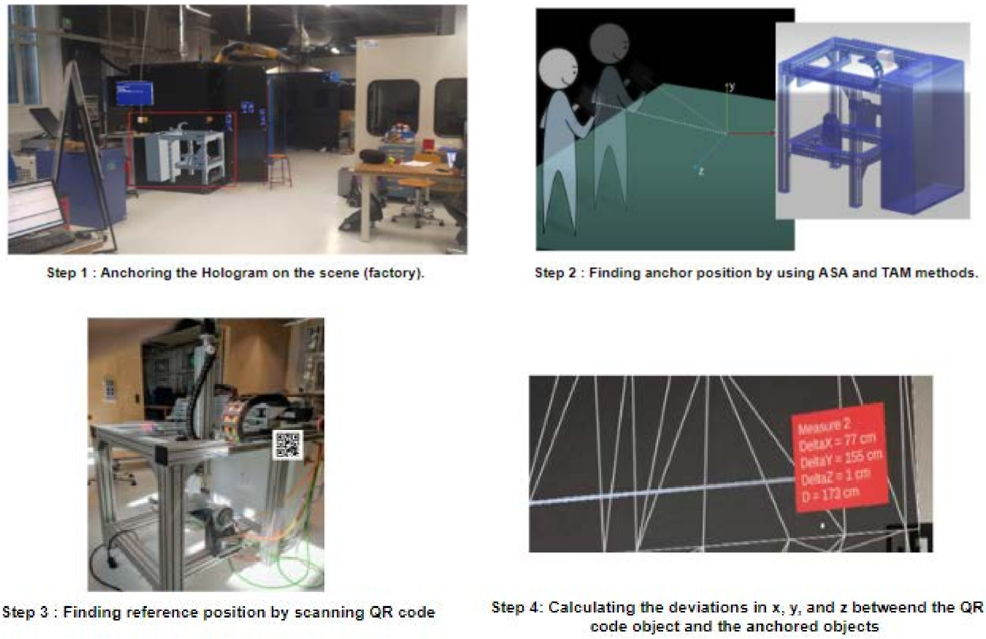


Figure 10. *Digital twin geometric coherence - deviation detection*

It's important to mention that an inventive hybrid approach (ASA+ Vuforia) by Wennan He et al. (2021) combines QR codes and Azure Spatial Anchors, but doesn't assess hologram positioning accuracy.

Results and Discussion

We compared both methods, evaluating precision, stability, network dependency, and data size. Precision was assessed by measuring the distance between QR code objects' nominal positions and anchored objects using both methods, with five measurements per anchor. The results are shown in the table below.

Table 1. *Precision - Deviation results*

Distance between QR code object and Anchored one	ASA method	TAM method
X-axis (Mean ± Standard Deviation)	4.568 ± 1.89 mm	3.82 ± 0.81 mm
Y axis (Mean ± Standard Deviation)	2.224 ± 2.1 mm	2.02 ± 0.93 mm
Z axis (Mean ± Standard Deviation)	3.512 ± 0.97 mm	4.718 ± 0.91 mm

Based on the results, it can be observed that the ASA method demonstrates greater consistency with lower standard deviations across all three axes. However, when it comes to precision, the advantage of one method over the other varies depending on the specific axis, but overall, the precision levels remain relatively consistent. The variability observed in the TAM method can be attributed to the quality of the wall detection algorithm. Therefore, we are actively seeking to enhance our technique to achieve greater precision and consistency in our measurements.

During the measurement process, we noticed that when we approached a confined area without capturing a significant portion of the scene, the yellow cube representing the ASA anchored object exhibited instability and began to exhibit slight movements (Figure 11).

Data size is crucial for improving digital twin coherence. TAM has minimal 5KB local data, while Azure requires an 80MB download. Tests yielded consistent results across internet speeds, with data size dependent on scene complexity.

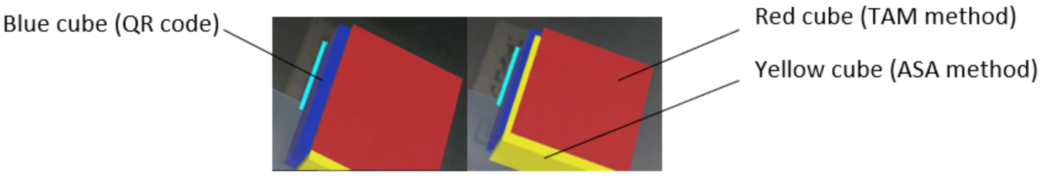


Figure 11. *ASA Object (yellow) Position Comparison: Large Scene Capture(right) vs Small Scene Capture (left).*

Contribution and Conclusion

In summary, our study emphasizes the TAM method's distinct characteristics, including local operation that reduces network dependence. While the ASA method excels in maintaining consistent precision along some axes, it relies on network connectivity and can show instability in certain situations. TAM's compact data size enhances data management efficiency. These findings stress the need for method selection based on specific use cases, with TAM presenting a promising alternative, especially in scenarios prioritizing network reliability and local operation. The comparison is summarized in Table 2

Table 2. *Summary results*

	TAM method	ASA method
Precision	Better in Z axis (depth dimension)	Better in X-axis
Positioning consistency	Higher standard deviation	Lower standard deviation
Stability	Stable Anchor	Anchor changes position
Network dependence	No (operates entirely locally)	Yes
Data size	Few kilobytes	Dozens of Mégabytes

From a scientific standpoint, our research aims to improve surface creation algorithms to address orientation challenges. This development can boost precision, consistency, and reduce task time, benefiting augmented reality applications. This aligns with the goal of enhancing AR system performance, promising a more robust digital twin positioning method.

References

Bo, J. Y. F., Shum, K., & Hong, K. Global Alignment of Meshes for the Microsoft HoloLens.
He, W., Xi, M., Gardner, H., Swift, B., & Adcock, M. 2021. Spatial anchor based indoor asset tracking. In 2021 IEEE Virtual Reality and 3D User Interfaces (VR), pp. 255-259

XR Application for Construction Progress Monitoring using 5G and BIM

Urs Riedlinger[1], Fabian Büntig[2], Marvin Voß[2], Mayra Fahrer[3], Tamara Graovac[3], Brian Klusmann[4], Duc Pham[4], Jessica Steinjan[5], Jan-Derrick Braun[5], Christian Geiger[2], and Leif Oppermann[1]

[1]Fraunhofer FIT, Sankt Augustin, Germany
[2]Düsseldorf University of Applied Sciences, Germany
[3]HHVISION, Köln, Germany
[4]eTASK Immobilien Software GmbH, Köln, Germany
[5]HOCHTIEF ViCon, Essen, Germany

Corresponding author: urs.riedlinger@fit.fraunhofer.de

Keywords: eXtended Reality, Building Information Modeling, 5G.

Introduction

This application integrates Building Information Modeling (BIM) with interactive eXtended Reality (XR) visualizations, leveraging 5G, to enhance construction. Our campus features a 5G stand-alone network covering indoor, outdoor, and currently a construction area. 5G's speed and latency improvements facilitate collaboration using BIM data and expedite data transfer for larger datasets, opening new possibilities at the site. Use cases encompass construction progress monitoring, orientation, warranty management, process visualization, and in-situ project display. We've developed a prototype enabling model manipulation on a tablet, alignment with surroundings, transparency adjustments, and process navigation, alongside site scheduling integration. Key research questions target the construction industry's support, public engagement benefits, and preferred interaction techniques for construction data. Our project aims to uncover 5G, XR, and BIM potentials for construction, highlighting advantages and challenges through user-centered iterations.

Related Work

Our paper focuses on three main subjects: XR, BIM, and interactions. BIM visualization in XR dates back to Sutherland's 1968 "sword of Damocles," the earliest head-mounted display (HMD) system outlining buildings (Sutherland, 1968). Recent instances include Raimbaud et al.'s off-site construction supervision with XR, BIM, and drones (Raimbaud et al., 2019). Chu et al. highlight mobile BIM-AR potential with QR codes (Chu et al., 2018), while Etzold et al. present web-based construction planning (Etzold et al., 2014). Sanchez-Sepulveda et al. emphasize XR in AEC education (Sanchez-Sepulveda et al., 2019), and Boos et al. explore AR for public participation (Boos et al., 2022). Our tool leverages XR and BIM to aid on-site construction management. Khairadeen Ali et al. enable remote monitoring (Khairadeen Ali et al., 2021), and Lin et al. link 4D BIM to construction using markerless tracking (Lin et al., 2019). AR interactions lack standardization, but Google and Apple offer useful interface guidelines (Apple, n.d.; Google, n.d.). For BIM data in Unity apps, Bille et al.'s workflows from 2014 remain pertinent (Bille et al., 2014). We adapt their approach with newer conversions and libraries.

Use Case Development

In the prototype development, we initially focused on implementing the construction progress monitoring use case. This involves utilizing the 3D BIM model to oversee project progress within the framework of project control. On-site users can compare the real-time construction status with the planned progress based on the project schedule. The AR application presents the 3D BIM model, synchronized with the construction timeline, allowing a direct comparison between the two states. Deviations from the planned timeline are visually highlighted, aiding in swift identification of disparities. Data captured on-site through the AR app is fed back into the scheduling software for prompt updates, enhancing the detection of delays and deviations. Advantages over the current approach include direct data capture via the provided 4D model (3D BIM and timeline) during site inspections, eliminating the need for subsequent data documentation and transfer. This accelerates schedule updates by directly conveying entered data to relevant participants, facilitating rapid analysis of deviations and enabling quicker, targeted countermeasures. Challenges involve ensuring the presence of a 3D model linked to a timeline for augmented on-site visualization. By placing markers at key points during a preliminary site inspection, serving as virtual anchors in the 3D model to align it with reality, it is possible to address this.

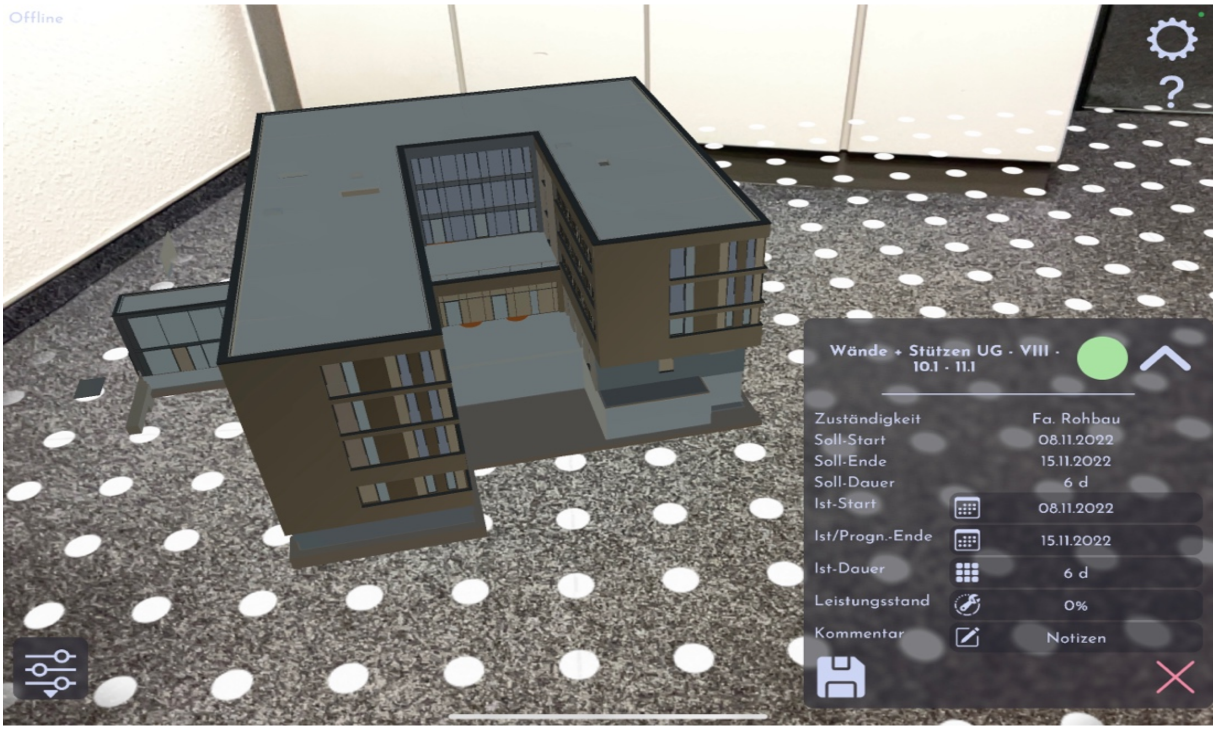


Figure 12. Screenshot of our prototype in the office. It shows the building in a smaller scale placed on the floor. On the right one can see an entry of the schedule with options to report the progress.

Figure 13. *Testing of our prototype on the campus' construction site.*

System and Interaction Design

The prototype (Figure 12, Figure 13) primarily loads data from a Common Data Environment (CDE), typical for BIM apps. 5G enables handling larger data sizes due to improved on-site bandwidth. Our data, derived from an Industry Foundation Classes (IFC) file, undergoes division: we extract the 3D model using IfcOpenShell-Toolkit, associate model data using a custom Unity parser, and integrate a construction process schedule. For warranty and defect management, we use BIM Collaboration Format (BCF) files. Our approach follows OpenBIM principles with standardized formats. Initially, a tablet app is deployed, catering to user familiarity. Interaction design based on common patterns and content relevance is prioritized. Incorporating Head-Mounted Displays (HMDs), like the Meta Quest Pro, for specialized scenarios is envisioned. Device categories encompass smartphones, tablets, HMDs, and stationary screens. Each category presents unique interaction rules, necessitating tailored UI approaches. Our software architecture prioritizes flexibility. Components follow Model-View-ViewModel (MVVM) patterns, split for reusable GUI views. Global stylesheets enable design flexibility. Mediator classes manage UI connections, ensuring the Model's independence from UI. Crucial Model scripts implement ScheduleTaskManager and IScheduleTaskVisualizer, handling schedule management and geometry visualization. To simplify comprehension, we use a LEGO model for demonstration (Figure 14). This choice aids in development, communication, and modular function building. Our LEGO dummy reflects the original with slight tracking adjustments. We leverage the AppCenter SDK for testing, ensuring up-to-date versions for testers and utilizing Unity's Cloud Diagnostics for remote issue tracking.

Discussion

Our system's development revealed distinct advantages and disadvantages, discussed next. Modular implementation with the LEGO test model facilitated swift idea adaptation and allowed early testing for valuable partner feedback. Mandatory updates ensured uniform testing with the latest version. Initial marker tracking and iPad LiDAR usage received positive

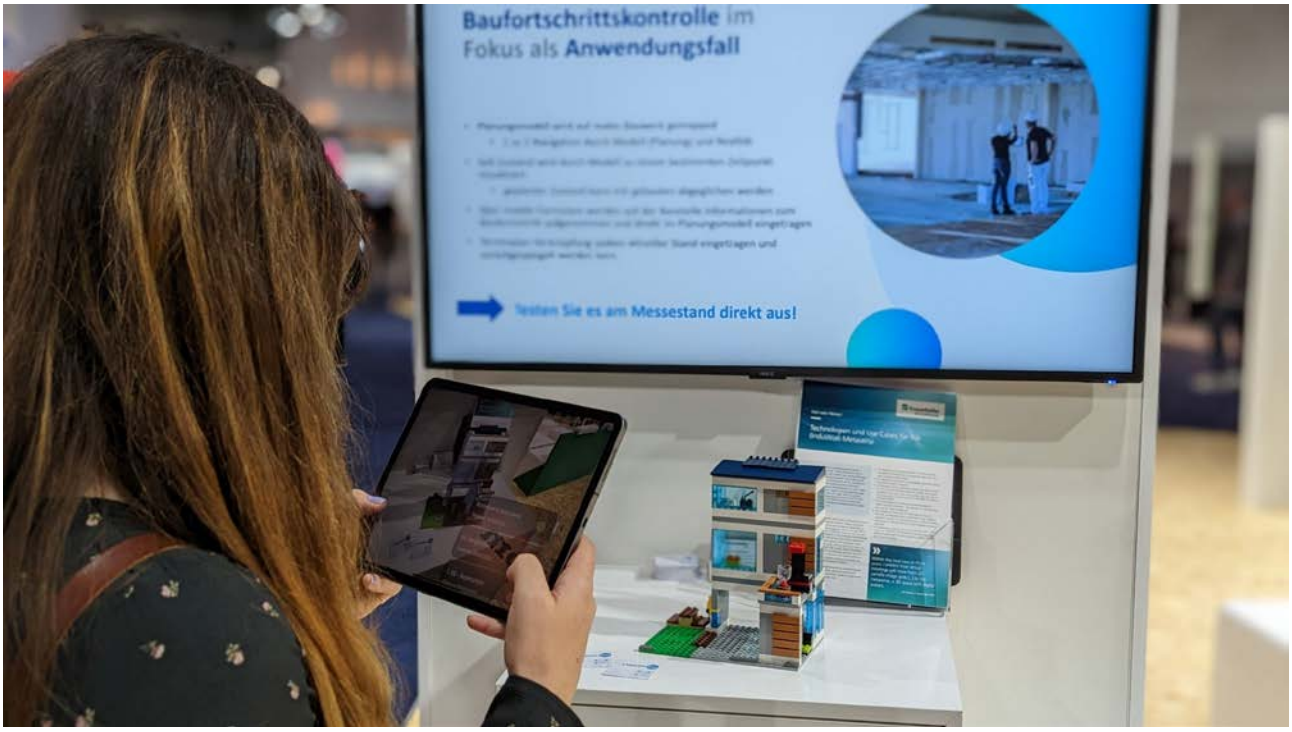


Figure 14. LEGO Model for demonstration purposes

feedback in user studies. GPS integration is planned for future navigation use cases, while remote collaboration features, avatar-based digital model interactions, and public presentations are being considered. Returning to our research questions, our prototype demonstrates XR and 5G benefits for construction. Careful use case consideration is crucial, as XR aids in comprehending plans and data. 5G supports timely data loading and remote collaboration, given reliable connections. XR might not always be necessary, with user judgment guiding required information presentation in specific contexts.

Conclusion and Outlook

Through the mentioned use cases, we aim to explore XR's potential with 5G and BIM. Prioritizing these use cases was crucial. Enhancing user experience demands meticulous interaction design. Our paper details our modular system, initial prototype (focused on construction progress monitoring), and our approach for various device categories. Observing user collaboration across these interfaces holds future interest. Remote rendering is another important topic to consider in the future, especially when it comes to large and detailed building models, where envisioned devices may come to their limits. Our future goal involves studying 5G's impact on XR BIM applications, benefiting stakeholders and the wider public.

References

Apple. (n.d.). 2022. Augmented Reality—Technologies—Human Interface Guidelines—Design—Apple Developer. Retrieved December 12, from https://developer.apple.com/design/human-interface-guidelines/technologies/augmented-reality

Bille, R., Smith, S. P., Maund, K., & Brewer, G. 2014. Extending Building Information Models into Game Engines. Proceedings of the 2014 Conference on Interactive Entertainment, 1–8. https://doi.org/10.1145/2677758.2677764

Boos, U. C., Reichenbacher, T., Kiefer, P., & Sailer, C. 2022. An augmented reality study for public participation in urban planning. Journal of Location Based Services, 0(0), 1–30. https://doi.org/10.1080/17489725.2022.2086309

Chu, M., Matthews, J., & Love, P. E. D. 2018. Integrating mobile Building Information Modelling and Augmented Reality systems: An experimental study. Automation in Construction, 85, 305–316. https://doi.org/10.1016/j.autcon.2017.10.032

Etzold, J., Grimm, P., Schweitzer, J., & Dörner, R. 2014. kARbon: A collaborative MR web application for communicationsupport in construction scenarios. Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing, 9–12. https://doi.org/10.1145/2556420.2556793

Google. (n.d.). Augmented reality design guidelines | ARCore. Google Developers. Retrieved December 12, 2022, from https://developers.google.com/ar/design

Khairadeen Ali, A., Lee, O. J., Lee, D., & Park, C. 2021. Remote Indoor Construction Progress Monitoring Using Extended Reality. Sustainability, 13(4), Article 4. https://doi.org/10.3390/su13042290

Lin, Z. Y., Petzold, F., & Ma, Z. L. 2019. A Real-Time 4D Augmented Reality System for Modular Construction Progress Monitoring—ProQuest. https://www.proquest.com/docview/2268537622?fromopenview=true&pq-origsite=gscholar&parentSessionId=MujmlbkFDlKAL4OclzlvVZW1eQ6Pqp0NGnUFDMklDXc%3D

Raimbaud, P., Lou, R., Merienne, F., Danglade, F., Figueroa, P., & Hernández, J. T. 2019. BIM-based Mixed Reality Application for Supervision of Construction. 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 1903–1907. https://doi.org/10.1109/VR.2019.8797784

Sanchez-Sepulveda, M. V., Marti-Audi, N., & Fonseca-Escudero, D. 2019. Visual Technologies for Urban Design Competences in Architecture Education. Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality, 726–731. https://doi.org/10.1145/3362789.3362822

Sutherland, I. E. 1968. A Head-mounted Three Dimensional Display. Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I, 757–764. https://doi.org/10.1145/1476589.1476686

Making the invisible visible for off-highway machinery by conveying extended reality technologies

Martijn Rooker[1,2], Michael Burmester[5], Gerald Fritz-Mayer[2], Clemens Arth[3], Sebastian Lorenz[4], Volker Waurich[4], Manuel Kulzer[5], Anastasia Sergeeva[6], Kaj Helin[7], Markku Pusenius[8], Benjamin Geslot[9], Daniel Röck[10], and Pekka Yli-Paunu[11]

[1]TTControl GmbH, Wien, Austria
[2]TTTech Computertechnik AG, Wien, Austria
[3]Technische Universität Graz, Austria
[4]Technische Universität Dresden, Germany
[5]Hochschule der Medien, Stuttgart, Germany
[6]Université du Luxembourg, The Netherlands
[7]VTT Technical Research Centre of Finland Ltd, Espoo/Tampere, Finland
[8]Creanex OY, Tampere, Fnland
[9]Haption SA, Soulgé-sur-Ouette, France
[10]Prinoth SPA, Vipiteno, Italy
[11]Cargotec Finland OY, Tampere, Finland

Corresponding author: martijn.rooker@tttech.com

Keywords: Extended Reality, Off-Highway, Mobile Machines, Co-Design, Human-Centred

Introduction

The off-highway domain contains every kind of transportation system and mobile machinery which does not operate on main roads (i.e., being excluded from all developments done in autonomous driving research), but metaphorically also stands for all kinds of machinery that falls out of standard focus of XR research, encompassing remote maintenance or assistance methods deployed to stationary production systems in factories (i.e., in one-fits-all audio-video telepresence systems that help to resolve failures). Many workplaces involving off-highway vehicles and machinery, such as in snow grooming, logistics or construction domain, are highly demanding and dangerous for e.g., their operators. Even brief moments of inattentiveness or misjudgement of distances, velocities or environmental conditions may lead to critical situations, putting human lives or the integrity of the machines, equipment, or cargo at risk.

XR technologies offer advanced capabilities regarding the utilization and accessibility of complex data, especially during machine operation. However, XR also has the potential to substantially challenge the operators, e.g., through new forms of interaction distraction or information overload (see e.g., Canito et al., 2020). Also, machine operators as highly trained users of expert systems tend to have very high expectations and requirements for human-machine interfaces (HMI). New interaction concepts must meet these expectations or most likely they won't be accepted. The enhancement of conventional HMI through multimodal information and interaction technologies such as VR/AR glasses, screen- and environment-projections, have been increasingly discussed in the last decades in the off-highway domain (see e.g., Palonen et al., 2017), but hasn't reached a viable level of applicability, usability, and an overall positive user experience.

The overall scope of the THEIA^{XR} project is to apply a user-technology focused research and development approach that focuses on the enhancement of the user-technology and the utilization of the potential of extending reality technologies and functionalities. The results will make the invisible visible respectively the non-perceivable perceivable to the human operator, extending the perceivable range of the operator without negatively influencing the performance of the human while controlling the machinery. The consortium consists out of 11 partners (universities, research institutes, SMEs and large enterprises) originating from 6 different countries. In three use cases in the field of (mobile) machinery from the off-highway and logistics domain, the implementation of novel eXtended Reality technologies and modalities will be investigated and developed.

Figure 15. THEIAXR logo

Technologies and Methodologies involved in the Development Process

Possible XR Technologies for Off-Highway Machinery

The application of XR technologies in the off-highway domain requires careful thoughts about the practical usefulness of visualization and sensing hardware available. Although developments in head-worn and hand-held devices have made significant progress over the last couple of years, the operation of off-highway machinery requires the full attention by the operator and usually both hands, rendering the use of handheld devices in their natural form implausible. Similarly, head-worn devices in non-remote operation scenarios are hardly applicable for reasons of wearing comfort, energy consumption or – most critically – safety concerns, not even mentioning device tracking problems in moving cabin scenarios and so on. Overall, XR technologies need to bring a significant benefit to the operator and aid, rather than obstruct the task to be fulfilled.

Considering the aforementioned conditions, the developments undertaken in THEIA^{XR} focus on technologies less traditionally used in XR. This relates to tracking, feedback, and sensing methodologies likewise, including, e.g., panoramic, and thermal imaging for data acquisition or laser and light-beam projection, amongst others.

In terms of visualizations, display-based interfaces are still valuable for presenting complex information. However, they often fail to integrate the presented information into actual task and attention flows. Therefore, we target the use of projections and multimodal-based XR in addition to display terminals.

Displays already play a central role in today's HMIs as they enable presentation of complex and dynamic data (e.g., machine state, environment surveillance, task and site management). However, these do not use augmenting visualization technics that utilize the current state of the art so far. However, displays, often mounted on the armrest or A-pillars are not ideal in workspaces that require intense observation of the environment most of the time.

Projection based XR directly places the virtual content in the world by light sources. To register the content accurately with respect to the off-highway machinery and the operator, respectively, the projection setups form rigid assemblies with the machinery itself, enabling the use of on-board telemetry data about position, velocity, or direction. This also facilitates the use of global information, such as surface models or georeferenced landmarks, which can be used to enrich the amount of information presented to the operator.

The availability of haptic rendering and/or force feedback is desired by most operators as it transmit direct feedback, especially important for precise control of movements or machine setups. The challenge is to offer a robust solution, which fits well into the confined space of a machine cabin, without completely breaking with current ergonomics. Of course, the hardware device must deliver signals far superior to the ambient noise (coming from sounds, vibrations, shocks, accelerations, inclinations, etc.). And it needs to make it possible to deliver diverse but coherent information, e.g., to avoid obstacles, to guide along optimal trajectories, or to attract towards desirable operating points in terms of energy consumption or impact to the environment.

Ambient light and acoustic feedback can support conventional ways of information presentation. Even though these modalities limit complexity of coded information, they can improve attention allocation and peripheric perception. Using our capabilities of mental sensor fusing offers the potential to allow cognitive processing of a much richer information environment for improved situation awareness. However, capabilities and limitations of multimodal interaction in machine operation contexts lack practical and empirical evidence on a sufficient level of technology integration.

With respect to the use of sensing technologies, recently developed methods available like e.g., the LIDAR sensor in Apple iPad Pro devices is of relatively little use, because the sensing range is limited, respectively the technology is not suitable to work plausibly well in outdoor conditions, like during snowfall or rainfall. What hasn't been investigated much in XR technology yet is the use of thermal imaging, which becomes particularly handy in scenarios with limited sight. Because the information given by such cameras is rather abstract, additional object detection technologies leveraging Deep Learning are employed to translate the information into plausible and assessable information for the operator. This refers to both the type of object, as well as its location relative to the machinery. Similar to the visualization methods employed, using a rigid assembly on the machinery enables the localization of said object with respect to the machinery and further reasoning about potential dangers and further actions.

Exploiting the XR prototype of the KALMAR's reach stacker for collecting user feedback of the remote operation

The main objective is to exploit the XR prototype of the KALMAR's reach stacker for the remote operation to collecting user feedback before actual implementation. The XR set-up was based on Varjo XR-3 which allows user to see the actual remote operation station and virtual representation of the reach stacker and harbour environment (see Figure 16). The user was controlling the reach stacker with joystick and keyboard and virtual environment with gestures.

The first user evaluation was held at VTT's XR lab with KALMAR personnel on 16.5.2023. There were four test subjects and they executed simple pick and place tasks. To assess the usability of the prototype, users were asked to fill in a System Usability Scale (SUS) (Brooke, 1996). The first prototype achieved a score of 59, which means usability was evaluated as marginal, but not acceptable, and following improvement was identified: (1) concept is working, but end-user is not willing to use HMD, (2) more augmented information to actual screens and XR environment, and (3) controls should in physical device e.g., in tablet or in joystick's buttons.

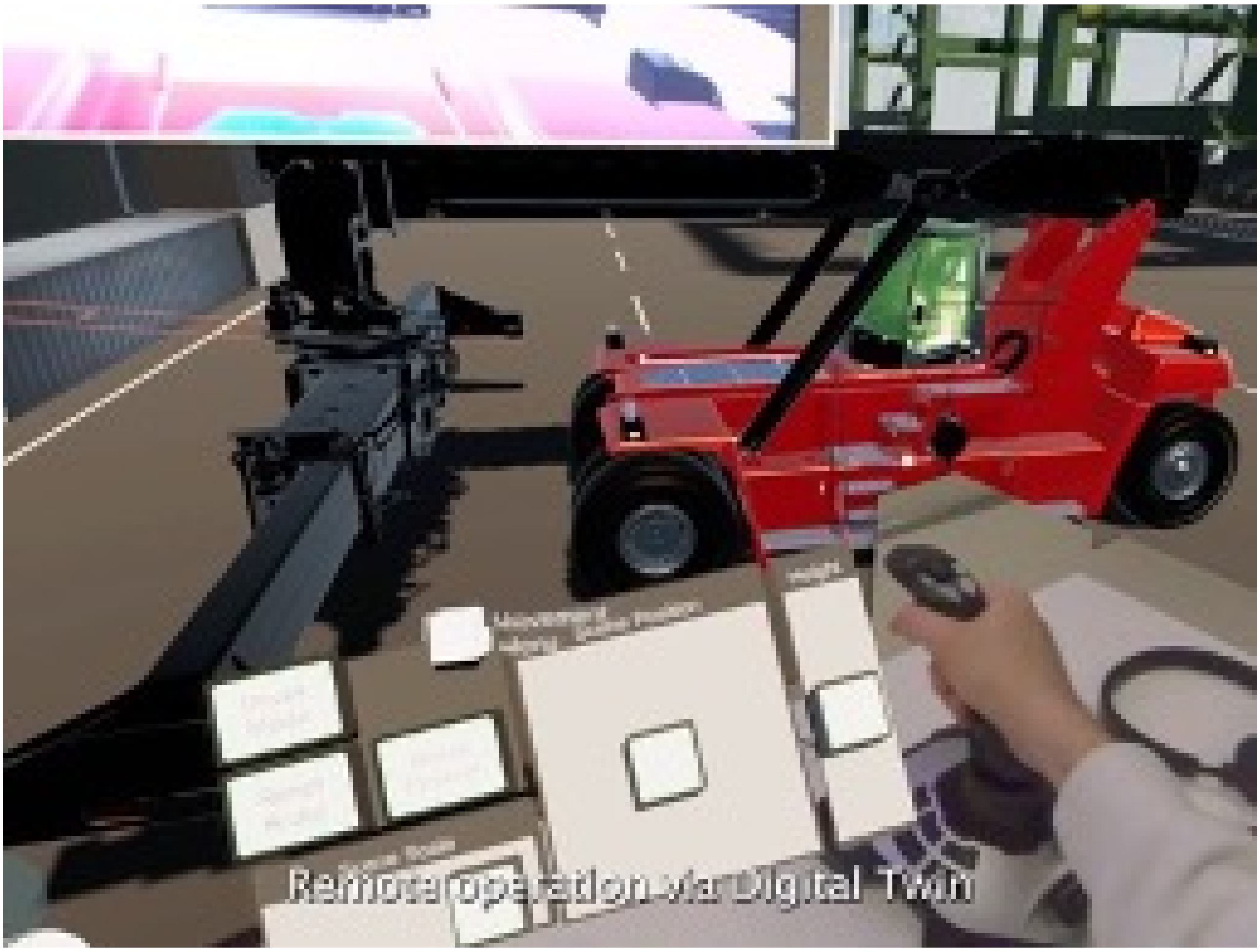


Figure 16. Mixed Reality based remote operation station which exploits Digital Twin (DT)

Transdisciplinary Co-Design

Interaction with technology impacts humans and their performances, attitudes, and well-being by design. Whereas the technological development of highly digitalized systems in all-off-highway domain progress, an accompanying design of operator workflows is often missing. It is the regarding of human states and capabilities that enable human-centred product development tools and methods to increase product quality of a solely technical discussion. Within our research, a core part of our XR human-centred methodology are concepts of usability (ISO 9241-11, 2018), user experience (Hazzenzahl, 2008), user-centred (Palonen et al., 2017) and conducive design Ziegler & Urbas, 2018), co-design and multimodal feedback (Schneidewind et al., 2016). The design approaches of human-centred design, conducive design and co-design will be combined in a transdisciplinary co-design methodology, involving all critical user- and stakeholder groups, scientific and technical disciplines as equal partners and experts in their respective fields. The methodology is depicted in Figure 17.

It is a four-phase approach, focusing in the first phase on the analysis of the context-of-use and involved tasks, followed by second phase, called co-design. Here a scenario-based design approach will be applied to create realistic usage scenarios for XR technologies in the respective use cases. Phase three is the experience prototyping, where a broad variety of prototypes with different fidelity levels will be used. The final phase is the co-evaluation involving end-users and stakeholders, testing, and validating the prototypes.

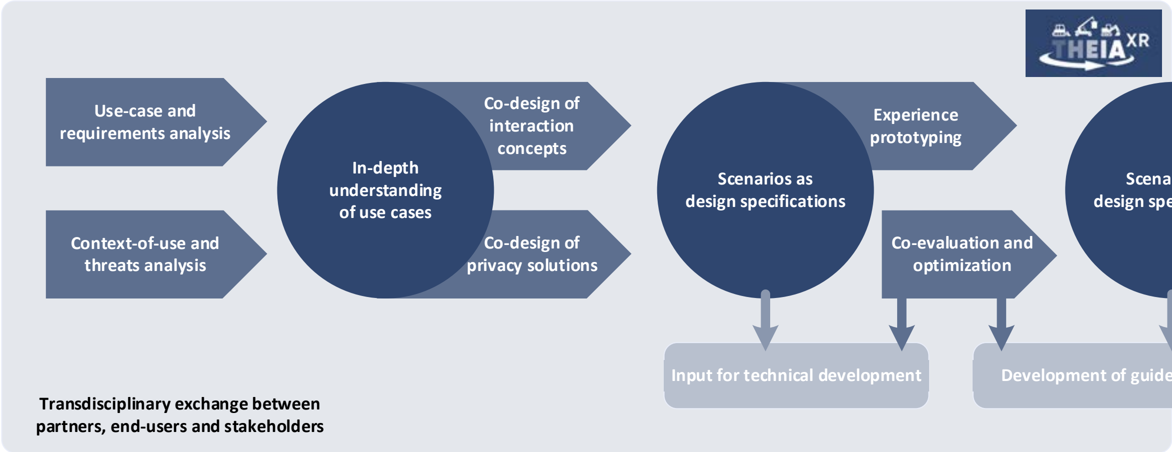


Figure 17. Visualization of transdisciplinary co-design methodology.

Addressing privacy concerns in extended reality for off-highway machinery

The development of XR solutions tailored to the off-highway domain brings forth distinct requirements concerning data privacy. Drawing from General Data Protection Regulation (2016) and the ISO/IEC 29100 Privacy Framework (ISO/IEC29100) as well as applied privacy-by-design methodologies (Pattakou et al., 2018), XR auditability framework, conceptualized by Norval et.al (2023) alongside works by de Guzman et al. (2019) and the XR Safety Industrial Initiative (2020), we've crafted a procedure to extract privacy risks grounded in expert analysis. The procedure involves an initial analysis of both the current workflow and the XR-enhanced workflow, as provided by experts. Each XR solution within the enhanced workflow is then examined for potential privacy risks, based on the Privacy Principles definitions and the groups of considerations outlined by Norval et al (2023). Experts then rate these solutions on a scale of 0 to 5, indicating their perceived likelihood of privacy risks posed by the technology. After this preliminary rating, concerns will be further discussed by a focus group of experts to determining the possible solutions to mitigate the risks. Building on research on creating user-friendly privacy and security features in both XR and Cyber-Physical environments (Abraham et al., 2022; Hanish et al., 2021; Oppl & Stary, 2022; Warinand & Reinhardt, 2022), and studies on the factors influencing general privacy perceptions (De Guzman et al., 2019, Coopamootoo & Groß, 2014; Xu et al., 2011), as well as those specifically in XR scenarios (Adams et al., 2018; Giaretta, 2022), we also pinpointed two key considerations for designing user-centred XR solutions for off-highway machinery: the users' perception of the utility of shared data and the importance they place on data privacy. Stemming from these insights, we devised semi-structured interview guidelines to gauge the current perceptions of end-users. These interview outcomes will serve as a baseline, enabling us to compare and contrast our solutions with existing ones regarding data control and privacy, setting the stage for collaboratively designing privacy-preserving mechanisms tailored for XR-enhanced operations. The proposed methodology will aid in seamlessly integrating XR solutions into existing industrial workflows without introducing additional privacy risks. It also can help making the privacy audit of existing industrial XR solutions in terms of potential privacy violations.

Use Cases

In the following, we are going to shortly describe the use cases and selected off-highway machinery representing them within the project. While their application domains are intentionally very different, they do have many things in common. Above all, they are operated by human operators, operating in low until no sight scenarios.

Snow Grooming Use Case

Figure 18. *PRINOTH snow groomer in operation*

Operation of snow groomers is becoming more and more difficult, as constantly more information is available which the operator of the snow groomer must take care of. Additionally, the environment around the machine is extremely difficult to handle (e.g., skiers around the machine, environmental conditions, etc.). Snowstorms causing no-sight conditions are a common situation that will occur, in which nevertheless the human operator must be able to handle the machine. Additional information provided using XR technologies can support the operator in guidance and operation, even in bad view of challenging conditions. One of the key problems in operating such a machine is that the human operator must constantly keep his hands on the controls of the machine and can thus not interact with the available data. The main objective here will be to supply the operator with needed information in the best possible and digestive way and to find a solution to drive with sensor and visualization assistance in a complete "white out" (no view to the environment possible).

Construction Use Case

Automation and digitalization are the most important development trends in the construction domain. Various assistance systems such as model-based excavation, automatic surveying, collision avoidance systems and tool management are state-of-the-art to increase productivity, quality, and safety of work. Despite their outstanding benefit for the operator, these systems are usually add-on-systems which are not integrated seamlessly in the HMI. Hence, they require additional multitasking skills to operate all relevant functionalities in parallel, which can lead to operators being overwhelmed and refusing to integrate these systems in their workflows. The THEIA^{XR} approach aims at providing a holistic XR-based HMI for the operator of an excavator. An intuitive and integrated HMI supports the operator in using digital assistance systems such as 3D-machine control, model-based earthworks, and collision avoidance. With the help of LiDAR-sensors, cameras, augmented video streams and ambient light, the various assistance systems are combined to a single, highly integrated HMI-concept.

Figure 19. Excavator in operation

Logistics Use Case

Within the logistics use case, the focus will be on (remotely) controlling a reach stacker (Figure 20), while handling containers in a harbour. In the current setup, operators operate with their eyes on video screens and remotely control these machines. From the footage, they observe how the container approaches the landing point on e.g., a truck or another pile of containers and how it settles in place. Through such footage, it is extremely difficult to perceive depth, which is a must when positioning the container. Supporting the operator to perceive three dimensionally and stay aware of items not visible in the cameras is a challenge when the machine must operate longer distance remotely. Safe operation requires that operators have enough information about surroundings. Since the visibility and feeling of presence (sounds, vibration, and movement of the machine) are not going to be on the same level in remote control operation than driving a machine from the cabin. It is very important that the operators feel that they are controlling the machine in a safe and predictable manner, and they have enough information about the surroundings.

Figure 20. KALMAR reach stacker in operation

Conclusions and Future Work

The usage of extended reality technologies has already shown great benefits in various domains. In the off-highway domain with mobile machinery, XR technologies are not deployed in a large fashion and still leaves many open research topics. Within the THEIA^{XR} project, the consortium wants to investigate the usage of XR technologies and investigate how the human operator can benefit. It will deploy a user-based co-design approach, which will involve human operators in all stages of the project, that is the requirements, design, implementation, and validation phase, to identify the problem areas and investigate how XR technologies can support the operator. The project is still at an early stage and started identifying the requirements for the different use cases and interacting with end users. In the next phase, first prototypes of XR technologies will be developed based on the feedback from the operators. The three use case domains (snow grooming, logistics and construction) have shown great interest and potential for these technologies and will proof to be challenging domains integrating these technologies, which will be done in the upcoming phase of the project.

Acknowledgements

This project has received funding from the European Union Horizon Europe research and innovation programme under grant agreement No 101092861. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.

Funded by the European Union

References

Abraham, M., Saeghe, P., Mcgill, M., & Khamis, M. 2022. Implications of XR on Privacy, Security and Behaviour: Insights from Experts. In Nordic Human-Computer Interaction Conference. 1–12. https://doi.org/10.1145/3546155.3546691

Adams, D., Bah, A., Barwulor, C., Musaby, N., Pitkin, K., & Redmiles, E. 2018. Ethics emerging: the story of privacy and security perceptions in virtual reality. In SOUPS@ USENIX Security Symposium. 427–442. https://www.usenix.org/conference/soups2018/presentation/adams

Brooke, J. 1996. Sus: a "quick and dirty'usability. Usability evaluation in industry, 189(3), 189-194.

Canito, A., Mota, D., Marreiros, G., Crchado, J.M., & Martins, C. 2020. Contextual adaptative interfaces for industry 4.0. In International Symposium on Distributed Computing and Artificial Intelligence. pp. 149-157. Springer, Cham. https://doi.org/10.1007/978-3-030-53829-3_14

Coopamootoo, K., & Groß, T. 2014. Mental models for usable privacy: A position paper. In Human Aspects of Information Security, Privacy, and Trust. Proc 2nd Int. Conf. HAS, p. 410-421, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27. https://doi.org/10.1007/978-3-319-07620-1_36

De Guzman, J. A., Thilakarathna, K., & Seneviratne, A. (2019). Security and privacy approaches in mixed reality: A literature survey. ACM Computing Surveys (CSUR), 52(6), 1-37. https://doi.org/10.1145/3359626

European Commission. 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). https://eur-lex.europa.eu/eli/reg/2016/679/oj.

Giaretta, A. 2022. Security and Privacy in Virtual Reality–A Literature Survey. arXiv preprint arXiv:2205.00208. https://doi.org/10.48550/arXiv.2205.00208

Hanisch, S., Arias-Cabarcos, P., Parra-Arnau, J., & Strufe, T. 2021. Privacy-protecting techniques for behavioral data: A survey. arXiv preprint arXiv:2109.04120. https://doi.org/10.48550/arXiv.2109.04120

Hassenzahl, M. 2008. User experience (UX): towards an experiential perspective on product quality. In Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine, 11–15. Retrieved from http://portal.acm.org/citation.cfm?id=1512717.

ISO 9241-11. 2018. Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts.

ISO/IEC 29100:2011. 2011. Informationtechnology Securitytechniques-Privacyframework Standard International Organization for Standardization, Geneva, CH. https://www.iso.org/standard/45123.html

Norval, C., Cloete, R., & Singh, J. 2023. Navigating the Audit Landscape: A Framework for Developing Transparent and Auditable XR. Proc. ACM Conference on Fairness, Accountability, and Transparency. 1418–1431. https://doi.org/10.1145/3593013.3594090

Oppl, S., & Stary, C. 202). Motivating Users to Manage Privacy Concerns in Cyber-Physical Settings—A Design Science Approach Considering Self-Determination Theory. Sustainability 14, 2, 900. https://doi.org/10.3390/su14020900

Palonen, T., Hyyti, H., & Visala, A. 2017. Augmented Reality in Forest Machine Cabin. Proceedings of the 20th World Congress of the International Federation of Automatic Control (IFAC 2017). IFAC-PapersOnLine. 50 (1), 5410-5417. ISSN 2405-8963 (electronic). https://doi.org/10.1016/j.ifacol.2017.08.1075.

Pattakou, A., Mavroeidi, A.-G., Diamantopoulou, V., Kalloniatis, C., & Gritzalis, S. 2018. Towards the design of usable privacy by design methodologies. Proc. IEEE 5th International Workshop on Evolving Security & Privacy Requirements Engineering (ESPRE), 1–8. https://doi.org/10.1109/ESPRE.2018.00007

Schneidewind, U., Singer-Brodowski, M., Augenstein, K., & Stelzer, F. 2016. Pledge for a transformative science: a conceptual framework. Wuppertal papers: 28. Retrieved from http://nbn-resolving.de/urn:nbn:de:bsz:wup4-opus-64142.

Warinand, C., & Reinhardt, D. 2022. Vision: Usable Privacy for XR in the Era of the Metaverse. Proc. European Symposium on Usable Security, 111–116. https://doi.org/10.1145/3549015.3554212

XR Safety Initiative. 2020. The XRSI privacy framework. https://xrsi.org/wp-content/uploads/2020/09/XRSI-Privacy-Framework-v1_002.pdf

Xu, H., Dinev, T., Smith, J., & Hart, P. 2011. Information privacy concerns: Linking individual perceptions with institutional privacy assurances. Journal of the Association for Information Systems 12 (12), 1. https://doi.org/10.17705/1jais.00281

Ziegler, J., & Urbas, L. 2015. Förderliches Gestalten komplexer Mensch-Maschine-Systeme: Eine disziplinenübergreifende Herausforderung. In: GfA e.V., ed. Arbeitswissenschaft. Mit Interdisziplinarität und Methodenvielfalt. Dortmund: GfA-Press.

Enabling X-Ray Vision via Multi-Camera Fusion for AR Applications

Ioannis Pastaltzidis, Iason Evaggelos Karakostas, Nikolaos Dimitriou, and Dimitroios Tzovaras

Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki, Greece

Corresponding authors: gpastal@iti.gr, iason@iti.gr, nikdim@iti.gr, dimitrios.tzovaras@iti.gr

Keywords: Augmented Reality, X-Ray Vision, 3D Pose Estimation

Introduction

While Augmented Reality (AR) enriches our Field-of-View (FoV) with virtual objects and annotations, it can also be used to enhance the users' situational awareness and perception (Apostolakis et. al, 2022). A useful application of AR is to make users aware of completely occluded objects. This work aims to address this type of problem, i.e. detecting occluded occupants of a building, using RGB(D) cameras and human 3D pose estimation towards annotating them on AR glasses. Estimating the 3D pose from a multi-view camera setup is a topic that has attracted a lot of attention in recent years. Multi-view, multi person 3D pose estimation in general lacked in speed when compared to single-view methods. The method introduced by Ye et. al (2022) has made great strides towards real-time multi-view pose estimation and is exploited in the proposed pipeline. The projection of totally occluded 3D poses on AR glasses, enables a form of X-ray vision, since the AR user can see fully occluded occupants inside the building even under motion. A short video demonstration of the pipeline is provided in this link. This application targets first responders (police, paramedics, security guards) as its users and can be deployed in facilities that have CCTV system (malls, airports, etc.). The proposed system is designed for scenarios where obstacles obstruct the line of sight, making it challenging for users to observe what is behind them. It proves particularly valuable in situations where there's a critical need for real-time awareness of the posture and location of concealed individuals. For instance, in law enforcement during a security incident, police officers require precise information about the posture and position of room occupants before taking action. Similarly, such spatial information would be very useful for first responders, like firefighters, when entering spaces during emergencies.

X-ray Pipeline

The presented pipeline utilizes 2 triplets of calibrated cameras in two rooms separated by a solid, opaque object, in our case a wall. Both triplets of cameras have the same origin point, and each captured space has dimensions of 4.5x4.5x2 (X, Y, Z) meters. The 3D pose estimation method predicts the 3D poses, which are later passed through a spatial 3D tracker, where the detections are matched using the center of mass from the 3D joints. A person is matched in consecutive detections if the Euclidean distance to the previous center is less than 0.3m. We also utilize a 2D Siamese tracker (Zhu et. al, 2018) onto the projected bounding boxes from one camera to track people of interest, e.g., the AR glasses user, with higher accuracy. The 2D tracker in essence improves the accuracy of the 3D tracker. In Figure 21, we showcase the 3D pose estimation pipeline for the two rooms. As illustrated, an intermediate layer produces a message containing information about the 3D poses, the tracking IDs and affiliations, which is transmitted to the AR devices via a message broker.

With regards to the projection of the 3D poses onto the AR glasses, we utilize an Inertial Measurement Unit (IMU) to determine the orientation of the head. The initial yaw value cannot be extracted from gyroscope and accelerometer sensors, and in this respect, we have experimented with CNNs for head pose estimation, April Tags and Plane Detection. Experimental results indicate superior performance of April Tags.

The initial yaw value and IMU data are used for the head pose initialization. We fuse data from the gyroscope and accelerometer with different weights and the rotation matrix $\boldsymbol{R}$ can be easily computed from the fused yaw, pitch, roll angles. The translation vector $\boldsymbol{t}$, is assumed to be the head position predicted from the 3D pose estimation method. Each AR device has a unique identifier, enabling it to retrieve the correct 3D skeleton from the predicted ones and obtain $\boldsymbol{t}$. The system can handle more than one AR users, and their affiliations are decided by the unique ID of the AR device they are wearing. The 3D poses can then be projected to AR glasses using the camera intrinsic parameters and $\boldsymbol{R}|\boldsymbol{t}$ matrix.

Experimental Results

In this section, a quantitative evaluation of specific components of the system is presented as well as qualitative results for the whole X-Ray pipeline.

Towards evaluating the accuracy of the 3D pose estimation in the wild, we used frame sequences captured by 3 calibrated cameras. We deployed state-of-the-art 2D pose estimation method on the synchronized frames from each camera and deployed triangulation from the OpenCV Structure from Motion (SfM) library, to obtain the ground truth 3D joints location, by exploiting the 2D joints from each view, the camera intrinsic parameters and the $\boldsymbol{R}|\boldsymbol{t}$ matrix. The evaluation was done for 90 frame triplets. As evaluation metric, the Mean Per Joint Positional Error (MPJPE) was calculated:

$$\boldsymbol{MPJPE} = \frac{\sum_i^n \|\hat{\boldsymbol{y}}_i - \boldsymbol{y}_i\|_2}{\boldsymbol{n}} \quad (\mathbf{1})$$

where $\hat{y}_i$ is the i-th joint ground truth, y_i the model prediction and n the number of joints. In our experiment the MPJPE was 83 mm while the same model when evaluated on the Panoptic dataset with 3 cameras has an error of 31 mm. Thus, there is a slight increase in the MPJPE for our setup, however the 8 cm of 3D error is considered acceptable for the presented application. Moreover, in the experiments we have conducted latency is more important than the 3D error and we manage to achieve real-time inference for the whole pipeline (60ms). In Figure 22 we provide the projection of 3D poses to the 3 cameras in our setup, as well as the visualization of 3D skeleton in space.

In order to measure the efficacy of the head pose estimation algorithms for the initial yaw value, we conducted an experiment, where we used an Intel RealSense camera mounted on a stationary tripod. We considered the wall point where the distance between the wall and the camera was at minimum to be the zero yaw angle in these experiments. We measured the mean absolute error for each Euler angle, in 19 different yaw positions and for each yaw position in 5 different pitch positions. In total 85 positions were measured. Results are presented in Table 3 and Table 4, with the first table referring to the April Tag pose estimation while the second one to Plane Detection. It is evident that pose estimation from April Tag is superior to the head pose prediction from Plane Detection. It should be noted that the roll angle cannot be estimated from the Plane Detection method and thus it is excluded. In Figure 23, the experimental setup is presented. We should mention that April Tag is superior to CNN based methods as well, as they have a mean absolute error (MAE) greater than 3.25 in the datasets they are trained on and 6-10 MAE at best, in more demanding datasets (Zhou et. al, 2023).

Table 3. Camera Pose error per angle April Tag

Angle	Mean	Median	Standard Deviation
Pitch	1.632	1.335	1.333
Yaw	1.568	1.24	1.082
Roll	3.591	2.707	3.171

Table 4. Camera Pose error per angle Plane Detection

Angle	Mean	Median	Standard Deviation
Pitch	19.64	3.807	29.122
Yaw	8.785	3.455	11.134

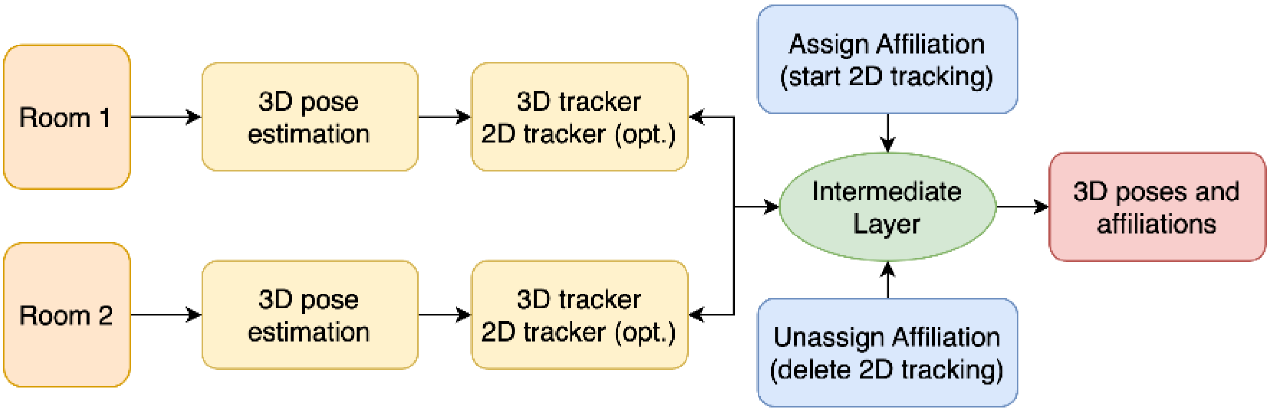


Figure 21. *3D pose estimation and tracking pipeline.*

Figure 22. *Projection of 3D poses and 3D visualization*

Figure 23. *Angle exp. setup*

Finally, we provide qualitative results for the projection of the 3D poses onto the AR glasses in Figure 24, where the outside space along with the detected person are visualized. In the left image in Figure 24, the red bounding box indicates the person encompassed within is being tracked, while the cyan bounding box is the bounding box from the projected joints. In Figure 25 we provide 2 views with the projected 3D poses back to the cameras while in Figure 26, we provide the AR visualization onto the glasses themselves. The projected poses in Figure 26 are correct and they are in-line with the projected poses in Figure 25. An additional example of AR visualization is provided in Figure 27.

Figure 24. Outside Room

Figure 25. Inside Room

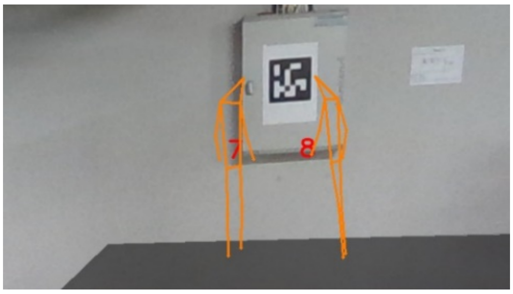


Figure 26. AR Visualization

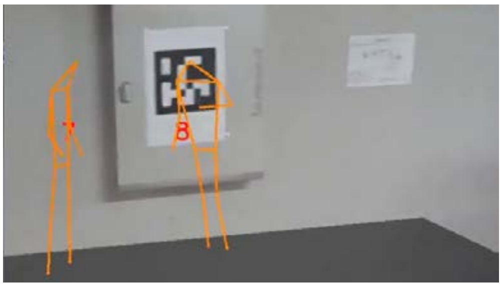


Figure 27. AR Visualization II

Acknowledgment

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883297 (project DARLENE).

References

Apostolakis, K. C., Dimitriou, N., Margetis, G., Ntoa, S., Tzovaras, D., & Stephanidis, C. 2022. DARLENE–Improving situational awareness of European law enforcement agents through a combination of augmented reality and artificial intelligence solutions. Open Research Europe, 1(87), 87. https://doi.org/10.12688/openreseurope.13715.2

Zhu, Z., Wang, Q., Li, B., Wu, W., Yan, J., & Hu, W. 2018. Distractor-aware siamese networks for visual object tracking. In Proceedings of the European conference on computer vision (ECCV) (pp. 101-117).

Zhou, H., Jiang, F., & Lu, H. 2023. DirectMHP: Direct 2D Multi-Person Head Pose Estimation with Full-range Angles. CoRR.

Ye, H., Zhu, W., Wang, C., Wu, R., & Wang, Y. 2022. Faster voxelpose: Real-time 3d human pose estimation by orthographic projection. Proc. Eur. Conf. Comp. Vision, pp. 142-159. Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-20068-7_9

A Framework for Assessing and Enhancing Presence in (a)symmetrical Remote Collaboration

Liv Ziegfeld[1], Maarten Michel[1], Ivo Stuldreher[2], Sylvie Dijkstra-Soudarissanane[3] and Omar Niamut[4]

[1]*Human Machine Teaming, TNO Soesterberg, the Netherlands*
[2]*Human Performance, TNO Soesterberg, the Netherlands*
[3]*Intelligent Imaging, TNO The Hague, the Netherlands*
[4]ICT, Strategy and Policy, TNO The Hague, the Netherlands

Corresponding author: liv.ziegfeld@tno.nl

Keywords: Social Presence, Spatial Presence, Quality of Experience, Computer-Mediated Communication, Remote Collaboration, Extended Reality, Expertise at a Distance

Problem Statement & Purpose

Extended reality (XR) technologies offer significant benefits for collaboration over more traditional forms of computer-mediated communication (CMC), by allowing to share richer media including visualizations and annotations, more non-verbal cues and increased sensory immersion. Hence, XR technologies are commonly employed in 'remote collaboration', in which individuals cooperate on a shared task and exchange knowledge while being geographically separated from one another. Despite these benefits, users of XR technologies for remote collaboration often encounter challenges, especially if the nature of the mediated interaction is asymmetrical, which is the starting point of our research.

Within remote collaboration, a use case class exists that we define as "Expertise at a Distance", in which a combination of user traits, context of use or the applied technology, is inherently different for users within the same remote collaboration. First, users are situated in diverse geographical locations, with different location characteristics. One user might be present in one physical task environment, while others participate from another remote location. Second, asymmetry in terms of knowledge and skills often exist, requiring clear transfer of knowledge and expertise from a more knowledgeable user to the less experienced user. Last, the communication technology used by the different users can differ greatly. While XR is experiencing growing adoption, e.g. for gaming and lifestyle, alternative mediums like phones and pc's frequently serve as the preferred method of communication. In these use cases, the asymmetrical traits of collaborating remotely between communication partners cause issues (Wang et al., 2021), and can affect the user's quality of experience (QoE) when not accounted for properly in the design and system set-up.

When reviewing existing work on presence and QoE in CMC (e.g. Oh et al., 2018; Grassini & Laumann, 2020) we identified several research gaps that motivate the present work. First, there is a lack of guidelines for improving the QoE of mixed medium collaboration, making it difficult to ground system design and application in past research. Second, there is a lack of publications that support hands-on assessment of QoE. While studies on measures for presence and QoE of (XR) communication systems exist, these often focus on comparing these measures (e.g. Grassini & Laumann, 2020), or on specific aspects of technology (e.g. Witmer & Singer, 1998) or the user (e.g. Bailey et al., 2016). Lastly, recent research on QoE for remote collaboration is centered around XR applications (e.g. Piumsomboon et al., 2019), while at this point in time, traditional CMC is still the norm in remote collaboration (e.g. Microsoft Teams, FaceTime), onto which XR is integrated or used as a replacement. Therefore, our framework should provide guidance for both existing and cutting edge CMC.

To tackle these research gaps, and to help developers and future users of CMC navigate through the complexities of their remote collaboration use cases, we developed a novel framework. This framework breaks the QoE down into detailed subcomponents, assists users in identifying the most important subcomponents for a scenario, lists how these factors can be measured using validated questionnaires, and lastly outlines baseline requirements and system enhancements, to evoke the desired level of QoE. The framework thus assists developers and users to design, specify and evaluate their CMC system for a remote collaboration use case.

Method

The present framework is based on the research conducted by Toet et al. (2022), wherein the significance of the presence construct is guaranteeing a satisfactory QoE (Skowronek et al., 2022; Toet et al., 2022). Toet et al. state that social- and spatial presence are both essential for allowing natural interactions and QoE in remote collaboration environments. In their Holistic Mediated Social Communication Questionnaire (H-MSC-Q), Toet et al. (2022) divide social and spatial presence into five abstract processing levels (cognitive, behavioral, emotional, sensory and reasoning). While these all support in assessing the level of presence, they provide little clue on where the system and its functions need change to increase the levels of presence. Therefore, while we consider the H-MSC-Q useful as a measurement tool, we focused on describing a more pragmatic method to achieve high levels of QoE.

Based on existing literature , we divided social presence and spatial presence into six subscales that directly relate to what a CMC system should be able to provide the user. As such, we divided social presence into the subscales affective interaction, intellectual interaction, and co-presence, whereas spatial presence consists of telepresence, ownership and agency (see Figure 28). Additionally, we identified more detailed factors with accompanying questionnaires from existing literature that fall under these six presence subscales. This understanding of the lower-level factors allows us to provide practical guidance on how to achieve the respective levels of presence.

After identifying the subcomponents and associated questionnaires, we established importance tests for each of the subcomponents, to promote reflection on whether a given subscale is of particular importance to the use case in question. We recommend that a baseline level of each presence subscale should be met to allow for an adequate level of QoE and thus effective communication. We define this baseline in terms of functional requirements in our framework and provide examples of related system requirements. For the presence subcomponents that are especially important for a given use case, an enhanced level should be strived for. Our framework delineates practical suggestions for how the requirements for the baseline level can be enhanced to allow for such an increased sense of presence. The full framework can be found on the Open Science Framework: https://osf.io/m36yp .

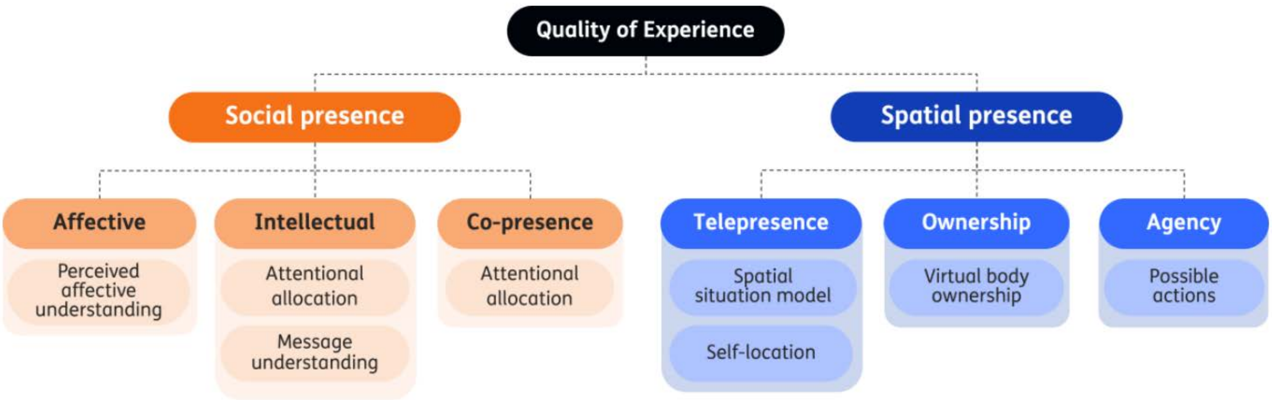


Figure 28. *Framework for presence in mediated communication based on the social and spatial presence subscales by Toet et al. (2022), with the respective subcomponents and measurements identified.*

Application to 'Expertise at a Distance' collaboration: Remote Maintenance

As an example of how our framework can be applied, we take a remote maintenance scenario. In this case a novice maintenance worker located on a wind farm (i.e. the physical task location) is confronted with some difficulties while repairing the gearbox of a turbine. They want to seek the assistance of a remotely located expert. The use case's asymmetrical properties and the need to form a detailed shared understanding of the task illustrates that our framework can also be applied to complex use cases. We investigate which forms of CMC should be employed, and in which manner, to allow for appropriate levels of QoE in the collaboration.

Firstly, the 'importance tests' related to each subcomponent from our framework can be used to identify which presence subscales are of greatest importance for the task at hand. While a baseline level of all six presence subscales should be ensured to allow for an adequate QoE, a few factors are of particular importance for this use case, namely intellectual interaction, telepresence and agency. While a high level of intellectual interaction is important for both the on-site worker and the remote expert for mutual understanding of instructions and questions, telepresence and agency are more important for the remote expert than for the on-site user. The remote expert should have an understanding of the worker's task environment (i.e. telepresence) and will likely want a degree of agency – i.e. the ability to control or show what needs to happen in the remote environment, to provide effective instructions. Equally, the remote expert will require a moderate to high level of telepresence to understand the spatial dimensions of the task environment to inspect the gearbox. The novice worker requires limited telepresence or agency of the other's virtual environment, since the to-be-performed task is entirely set in the novice's real surroundings.

When consulting our framework, we can also identify system recommendations. In this case, rich, synchronous backchannels are important to allow for an enhanced level of the subfactor of 'perceived message understanding'. This functional requirement could then be translated into the system requirement of including microphones and speakers in the set-up for both users, so they can easily communicate their level of understanding verbally and in real-time.

This approach can be used for all the different subscales, providing insight into what constitutes the subscales, how they can be measured, and what functional requirements are recommended to achieve adequate presence levels. The questionnaires corresponding to each subscale can then be used to measure the level of presence factors achieved, after which further system improvements can be made if necessary. Users working with an existing system can also employ these questionnaires. They can identify the subscales for their scenario and then use the questionnaires to evaluate whether their set-up leads to satisfactory outcomes on the presence subscales. It is important to remember that the requirements will differ for each scenario, but our framework can support identifying these unique properties. Further, one must consider the effects and hurdles when adopting new technologies, and therefore also when using our framework. Users will require time to get accustomed to new forms of CMC, and may have to adopt modified ways of performing their tasks with the new technologies.

Conclusions and Future Work

With our framework, we have created a prescriptive method to both assess and enhance the social and spatial presence of remote collaboration, making it easier for both researchers and developers to improve remote collaboration hands-on. As a next step, we aim to write a step-by-step guide on how to apply the framework and actually apply it to real world remote collaboration use cases, especially Expertise at a Distance use case class. Secondly, we want to research if and how the presence subscales are interrelated; do increased levels for each subscale always lead to better QoE, or are there inverse correlations? Lastly, we want to research the added value of objective measures of presence compared to subjective questionnaires.

References

Bailey, J. O., Bailenson, J. N., & Casasanto, D. 2016. When does virtual embodiment change our minds? Presence, 25(3), 222-233. https://doi.org/10.1162/PRES_a_00263

Grassini, S., & Laumann, K. 2020. Questionnaire Measures and Physiological Correlates of Presence: A Systematic Review. Front. Psychol. 11:349. https://doi.org/10.3389/fpsyg.2020.00349

Oh, C. S., Bailenson, J. N., & Welch, G. F. 2018. A systematic review of social presence: Definition, antecedents, and implications. Frontiers in Robotics and AI, 5, 114. https://doi.org/10.3389/frobt.2018.00114

Piumsomboon, T., Dey, A., Ens, B., Lee, G., & Billinghurst, M. 2019. The effects of sharing awareness cues in collaborative mixed reality. Front. Robot. AI 6. https://doi.org/10.3389/frobt.2019.00005.

Toet, A., Mioch, T., Gunkel, S. N. B., Niamut, O., and van Erp, J. B. F. 2022. Towards a Multiscale QoE Assessment of Mediated Social Communication. Qual. User Exp. 7, 4.

Skowronek, J., Raake, A., Berndtsson, G. H., Rummukainen, O. S., Usai, P., Gunkel, S. N., ... & Toet, A. (2022). Quality of experience in telemeetings and videoconferencing: a comprehensive survey. IEEE Access, 10, 63885-63931. https://doi.org/10.1109/ACCESS.2022.3176369.

Wang, P., Bai, X., Billinghurst, M., Zhang, S., Zhang, X., Wang, S., He, W., Yan, Y., & Ji, H. 2021. AR/MR remote collaboration on physical tasks: a review. Robot Comput Integr Manuf 72:1–32. https://doi.org/10.1016/j.rcim.2020.102071

Witmer, B.G., & Singer, M. J. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. In Presence, 7 (3), 225-240. https://doi.org/10.1162/105474698565686.

Building Industrial Metaverse: Technical Challenges of Real-World Collaborative XR Application Development

Yücel Uzun, Urs Riedlinger, Florian Buchholz, and Leif Oppermann

Fraunhofer FIT, Sankt Augustin, Germany

Corresponding author: yuecel.uzun@fit.fraunhofer.de

Keywords: Industrial Metaverse, Extended Reality, XR, Avatar, Digital Twin

Introduction

Regardless the ongoing discussions surrounding the concept of the metaverse, industrial potential of extended reality (XR) technologies has been researched and applied since 1960s (Sutherland, 1965) and various aspects of it has been investigated in controlled environments many times. With the rise of recent affordable but highly capable consumer XR devices, as well as improvements on the software environment surrounding them, the realization of industrial use cases has transformed from theoretical exploration in laboratories to practical applications in factories.

As part of the "IndustrieStadtpark" project, we have developed a remote maintenance and training system with our industry partners. In our earlier work (Oppermann et. al. 2023), we have outlined the project's objectives, provided an overview of its structure as well as the overall design, development process and how it stands compared to other systems. This paper shortly re-introduces the system, and then elaborates over the specific technical challenges we had to endure and overcome during the development process.

Overview of the Application

The main goal of the implemented solution is providing a seamless collaboration between the on-site technicians and the off-site experts during maintenance and training tasks for a real industrial setting. This is achieved by using digital twins, avatars and video-audio communication. Figure 29 shows the system in action.

The setup consists of 3 main parts:

1) the server infrastructure to communicate end devices and handle the data management

2) an XR application on HoloLens 2 for on-site technician and

3) a virtual reality (VR) application on Meta Quest devices through Quest Link system for the expert user. (Oppermann et. al. 2023) includes more comprehensive representation of the system setup.

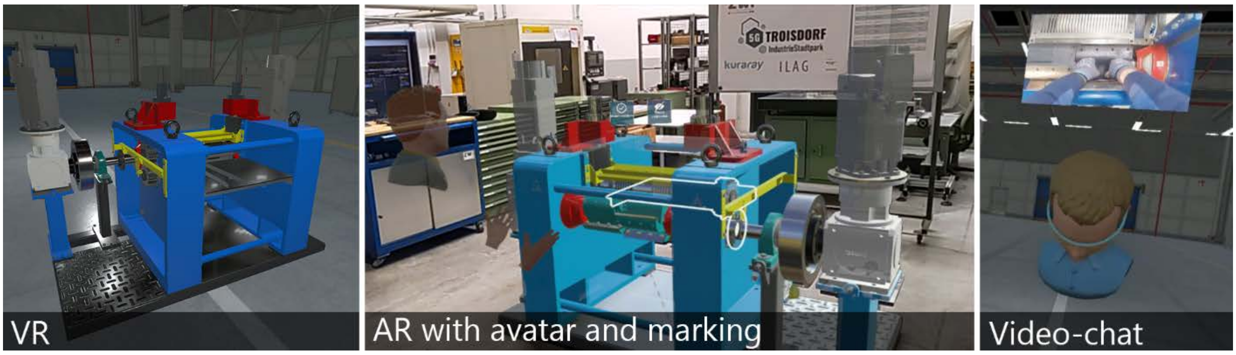


Figure 29. System overview. Experts view the digital twin, while technicians see an overlay on the real machine, allowing communication via avatars, video chat. They can also interact with the digital twin and add annotations.

Technical Challenges

Despite the recent improvements and standardization affords, such as OpenXR, XR is still a cutting-edge and marginal technology. Because of this, developing a real-world application requires dealing with various limitations and addressing small edge cases which sometimes create unexpected but significant issues. Moreover, unlike synthetic laboratory environments or entertainment use-cases, data privacy and security are serious concerns for the industry partners due to sensibility of the trademarked data.

The first and obvious issue we had to tackle was bringing digital twins to the limited-performance end-devices such as HoloLens 2. Digital twins of modern industrial machines, while available in our project, are not optimized for XR applications. Notably, they have extremely high details and, in some cases, even several screws of a model are enough to overload HoloLens 2. Manual reduction of polygon counts is a potential solution, but this approach is time-consuming and demands extensive domain expertise related to the real machine. Furthermore, automated mesh optimization tools like NuGraf or MeshLab are inefficient for these kinds of 3D-models since models have not been designed for these use cases. Considering our project's primary objective of building and evaluating a standalone 5G network, it became apparent that remote rendering was the most suitable approach.

When we have started the development, there were several frameworks that claimed to provide remote rendering functionality. However, our feasibility analysis showed only Azure Remote Rendering and Holo-Light had reliably working products that run on HoloLens 2, and some products like NVIDIA CloudXR had deprecated their HoloLens 2 support due to performance concerns. Given Microsoft Azure's pre-existing security clearance with our industry partners and it being conceived as more secure environment to host their trademarked digital twins compared to our own hosting, as well as the projected cost and easier worldwide server management reasons, we implemented Azure Remote Rendering in our system.

Unfortunately, at the time we have added the VR clients, Azure Remote Rendering lacked native support for Meta Quest platform. Because mobility was not a necessity for experts, we have compromised with utilizing local rendering on a PC via Quest Link system. This compromise, however, meant losing the hand, eye, and face tracking features of the hardware, capabilities we had intended to leverage for creating a more immersive and genuine communication environment.

Handling of the model caused another unexpected challenge during the development of the synchronization between clients. Specifically, users should be able to interact with specific parts of the model, which are defined hierarchically in the digital twin. Modern networking frameworks for Unity, such as Photon or Mirror use unique identifiers to distinguish and synchronize different objects over network, but this requires both client applications and the server to contain the same objects with the same identifiers. In contrast to this usual scenario in our case, only VR application had the complete object, HoloLens 2 client only had access to loose representation of the hierarchy through Azure services, and the server only hold the name of the active model. To overcome this, first we built our own identification system for Mirror based on the object names and their hierarchy, then created a state system that keeps track of the interacted parts.

Another issue that has been introduced by combining the synchronization system and remote rendering was the increase of information update latency on the HoloLens 2. For instance, if the expert marks a part of a digital twin, the HoloLens application first receives this information, then identifies the related parts through Azure queries, then makes another query to update the visualization of these parts. Even though these occur in millisecond levels under ideal conditions, in the real-world, networks and user interactions are unpredictable, and these cases are challenging to replicate in the development environment. Numerous edge cases in this retrospect had to be addressed to provide a consistent and reliable user experience while avoiding application terminations.

Unexpected application terminations and restarting the application posed another unique problem since we wanted to retain the sessional data until all tasks are completed. Unity engine, by default, defines the world origin as the start position of the app. Since the VR client does not have a direct connection with the real environment, we are able to assume the AR and VR coordinate systems align from the beginning. However, if the HoloLens application is restarted, this moves the origin of the application to another location in the real world. While we can still have the same assumption and update the location of the digital twin and avatar of the expert user seamlessly, this required us to store the any other spatial data relative, such as annotations, to the digital twin rather than origin, which resulted more complicated development path and required more intensive testing.

Positioning of the machine also created an unexpected prioritization conundrum. Placing the digital twin over the real object is an important feature: non-aligned placement might cause uneasiness on HoloLens user and create miscommunication with the expert. 3D object detection and tracking are well addressed problems and there are already commercial platforms that support HoloLens 2, such as Vuforia or VisionLib to tackle these issues. Due to performance reasons, these solutions are designed for ARM64 CPU architecture. However, video – audio communication is even more important feature, but options for it are more limited. As part of the Mixed-Reality Toolkit libraries, Microsoft developed and maintained a WebRTC implementation for HoloLens platform, but the development went idle before adding ARM64 architecture support and the project was eventually deprecated. However, as mentioned, existing object tracking libraries are designed solely for ARM64 architecture, and due to lack of better options, we used this deprecated but still working library and had to resort to employ the basic QR-Code detection and manual placement methods for placing the digital twin, which created sub-optimal user experience.

Discussion

Our experience with the XR development shows the clear distinction between the marketing hype and the reality. The software environment is not only riddled with fragmentation and incompatibilities, but also suffers from lack of resources and long-term support for existing frameworks. Nevertheless, these are understandable and natural due to dynamicity of the current state of the technology, and project management should account these factors while planning for a real case scenario. This requires identifying critical problems that the application aims to solve early on, running a feasibility analysis for these subjects and confirm if they are reliably and reasonably solvable with the current state of the software and hardware. It's important to stay flexible on non-critical or nice-to-have features and be ready to resolve these in different and creative ways.

XR and network-heavy use cases are also open to lots of edge-cases as mentioned in the previous chapter. This underscores the significance of adopting a user-centered iterative design and development process (ISO 9241-210). Based on our experiences, we advocate for early user involvement and frequent testing throughout the entire development cycle.

Although these challenges are inherent to software development in general, their impact is particularly evident in both commercial products and the private applications we have had the opportunity to test at various events. Despite the hardware improvements, it is worth noting that many industrial-metaverse applications remain relatively simplistic, primarily due to the factors previously mentioned and remain, although being stable, scalable versions of Google Glass application prototypes from the mid-2010s.

That being said, it is essential to acknowledge the significant progress made since the project has started, including, but not limited to, hardware developers embracing the OpenXR framework and software developers adopting standards like Universal Scene Description, the announcement of closer collaborations among influential companies such as Unity, Microsoft, and Meta, Microsoft's integration of Azure remote rendering support for Meta devices, and the transfer of the Mixed Reality Toolkit to a collaborative organization involving Magic Leap and Qualcomm. Furthermore, other standardization attempts like NVIDIA Omniverse and the entry of Apple into the XR specific hardware domain hold great potential. These developments increase the hope for enhanced standardization and stability within the XR software landscape, ultimately leading to more comprehensive and immersive user experiences.

Acknowledgments

This work was funded by the German Federal Ministry for Digital and Transport BMDV as part of the national 5x5G strategy. We thank our project partners from "IndustrieStadtpark".

References

Oppermann, L., Buchholz, F., & Uzun, Y. 2023. Industrial Metaverse: Supporting Remote Maintenance with Avatars and Digital Twins in Collaborative XR Environments. 2023 CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3544549.3585835

Sutherland, I. E. 1965. The ultimate display. In Proceedings of the IFIP Congress, 2, 506-508. Available at: http://www.universelle-automation.de/1965_Boston.pdf

Evaluation of a High-Fidelity Orthoptic Learning Simulation using Unreal MetaHumans

Georg Meyer[1], Jignasa Mehta[2], David Newsham[2], Ryan Ward[1], and Simon Campion[1]

[1]*Virtual Engineering Centre, University of Liverpool, United Kingdom*
[2]*School of Health Sciences, University of Liverpool, United Kingdom*

Corresponding author: georg@liv.ac.uk

Keywords: Simulation Fidelity, User Experience, Virtual Reality, MetaHumans, Medicine and Rehabilitation, Human Factors

Introduction

Orthoptics students develop extensive clinical diagnostic and psychomotor skills to characterize and treat eye movement defects and problems in binocular vision. Computer simulations of patients enable students to experience a wide range of conditions and acquire the necessary diagnostic skills in educationally suitable settings.

Key skills that students learn include:

1) Core diagnostic skills, for example the precise characterization of eye-movement limitations or deviations.
2) The necessary motor-skills to effectively use tools, such as a fixation targets (torch, fixation-stick), occluders, or prism bars to measure deviations.
3) Interpersonal skills, especially overcoming the reluctance to approach patients very closely for the examination.

The VR Simulation

Here we present a high-fidelity, freely configurable patient simulation using Unreal MetaHumans that enables students to develop key skills before they embark on placements in hospital settings where they interact with real patients (Figure 30 and Figure 31).

The use of a VR environment provides precise models for eye and lid movements as well as key information such as the corneal reflections or the optics of the prism bar. The hand controllers were configured for intuitive use: students pick up tool by briefly pressing the trigger button, after that the hand controller can be handled just like the tools, which, for example, teaches students to effectively swap tools between hands to examine both hemi-fields. The highly realistic rendering, finally, enables students to experience very close interaction with patients as well as the use of, for example, specular reflections in diagnostics.

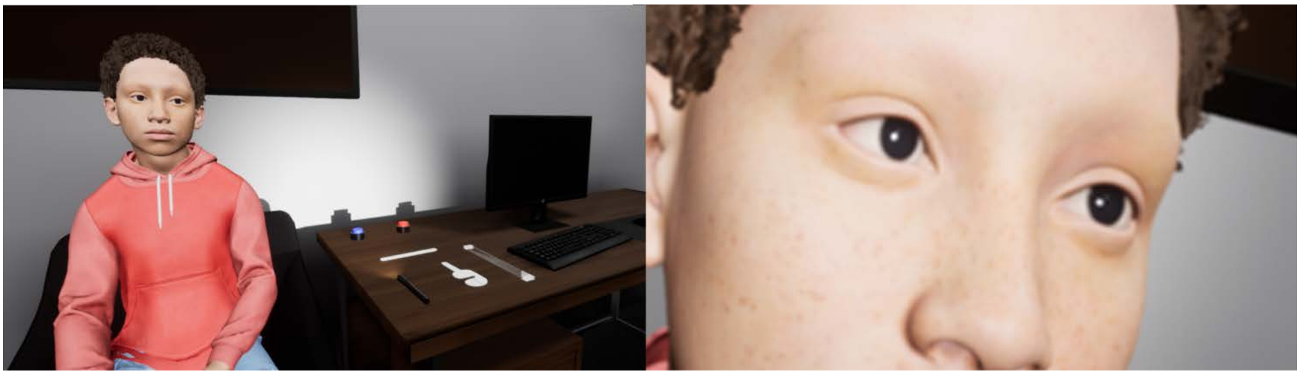


Figure 30. Users are presented with a virtual environment containing a freely configurable patient as well as a range of diagnostic tools. The use of MetaHumans provides a highly realistic and engaging visual environment. Details, such as the faithfully rendered corneal reflections are essential for the diagnosis of deviations and movement limitations.

The simulation was designed to maximise fidelity, eye and eye-lid movements, for example, were modelled realistically (Figure 31, left), the visibility of corneal reflections was optimised, and realistic lid movements were included.

Figure 31. Eye and lid movements as well as corneal reflections were modelled faithfully (left). Hand-controller use (right, here representing the torch) was designed to be as natural as possible: users press the trigger to 'pick up' any tool, but then are free t

Evaluation

This study compares user experience and simulation fidelity in the VR simulation with an existing screen-based educational simulation, the AAO complex strabismus simulator (Örge and Epley, n.d.). This web-based simulation presents a simplified 'cartoon face' with realistic eye movements, but, for example fixed corneal reflections represented by a white disk at the centre of the eye. The MetaHuman patients are rendered with great fidelity and in an immersive environment with realistic lid and eye movements as well as corneal reflections.

Two year-groups, year 1 and year 2 with 23 students each, evaluated both simulations using well-established scales: Simulation fidelity was measured using the seven-item modified presence questionnaire (mPQ), designed for use in clinical settings (Brackney & Priode, 2017). Both year groups rated fidelity of the VR simulation as significantly higher, providing the basis for our claim that we provide a high-fidelity simulation, Figure 32.

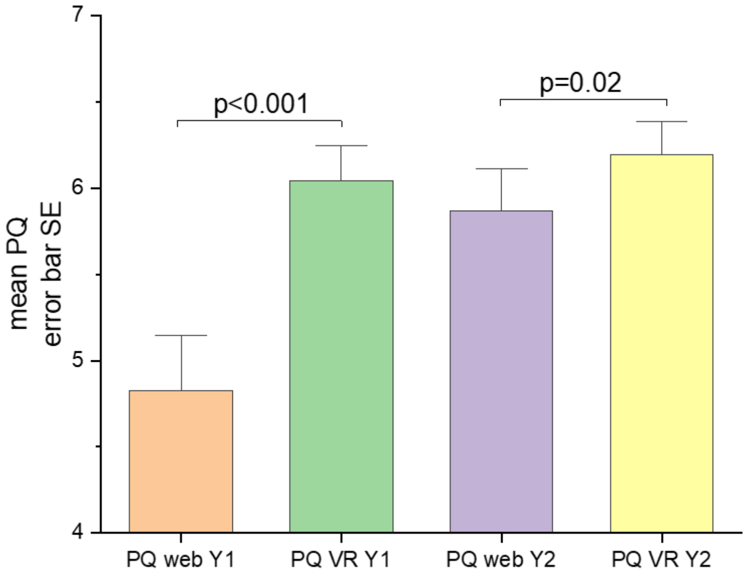


Figure 32. Users in both year groups (Y1 and Y2) used the mPQ scale to rate simulation fidelity for the VR simulation and a screen-based existing teaching tool on a seven-point Likert scale. The VR simulation scores significantly higher. The, more experienced, second year student group also rates the web-based AAO simulation (web) as highly realistic.

Subjective ratings for ***user experience*** were captured using the eight-item UEQ-S questionnaire (Schrepp et al., 2017), which provides separate measures for pragmatic and hedonic user experience, (Figure 33).

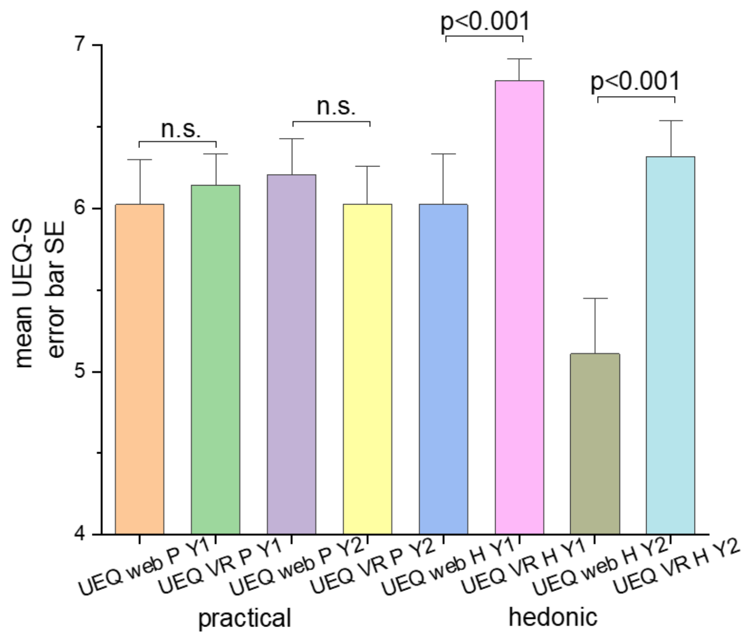


Figure 33. User-Experience, rated on a seven-point Likert scale, was measured using the UEQ-S questionnaire. The scale provides separate assessments of the practical and hedonic quality of the simulations. In both year groups (Y1 and Y2) we found no differences in the practical quality, but substantial improvement in the user ratings for hedonic quality in the VR simulation compared to the web solution.

There were no significant differences between the web-based and VR simulation in user experience for the *pragmatic* quality measure ($p > 0.05$). The VR simulation, however, scored significantly better in hedonic quality and simulation fidelity measures (both $p < 0.005$).

The latter two measures are highly correlated. To account for individual differences in the interpretation of the rating scales, scores for the VR and web implementation were subtracted from each other for each of the three (sub)scales. For both year groups, the hedonic UEQ-S subscale and simulation fidelity ratings were significantly correlated: Year 1, $r(22)=0.589$, $p = 0.003$; Year2 $p(22)=0.749$, $p=0.0003$.

Conclusion

Both simulations that were evaluated were designed to represent realistic eye movements in response to user actions. Key differences include user interaction (virtual tools vs mouse) and the rendering of the environment and face of the virtual patient. The observed significant improvement in simulation fidelity ratings therefore was a desired effect.

The mean user experience (UEQ-S) scores are excellent for both systems. It is notable that the practical ratings for the VR system are as good as those for the web-based interface.

Especially interesting is the finding that the hedonic user experience ratings are substantially better for the VR system. We expect these differences to lead to substantial differences in outcome measures, such as learning performance and learning transfer (Cooper et al., 2018, 2021). Future work should investigate the specific reason for these improvements.

We conclude that increased simulation fidelity can substantially enhance the perceived hedonic quality of a simulation learning package without affecting the perceived practical usability. Simulation fidelity is a key contributor to hedonic user experience.

References

Brackney, D. E., & Priode, K. 2017. Back to reality: the use of the presence question-naire for measurement of fidelity in simulation. Journal of Nursing Measurement, 25(2), 66E-73E. https://doi.org/10.1891/1061-3749.25.2.66

Cooper, N., Milella, F., Pinto, C., Cant, I., White, M., & Meyer, G. 2018. The effects of substitute multisensory feedback on task performance and the sense of presence in a virtual reality environment. PloS one, 13(2), e0191846. https://doi.org/10.1371/journal.pone.0191846

Cooper, N., Millela, F., Cant, I., White, M. D., & Meyer, G. 2021. Transfer of train-ing—Virtual reality training with augmented multisensory cues improves user experi-ence during training and task performance in the real world. PloS one, 16(3), e0248225. https://doi.org/10.1371/journal.pone.0248225

Meyer, G. F., Wong, L. T., Timson, E., Perfect, P., & White, M. D. 2012. Objective fidelity evaluation in multisensory virtual environments: auditory cue fidelity in flight simulation. https://doi.org/10.1371/journal.pone.0044381

Örge, F.H., & Epley, K.D. Complex Strabismus Simulator https://aao-resources-enformehosting.s3.amazonaws.com/resources/Pediatrics_Center/Complex-Strabismus-Simulator/index.htm l

Schrepp, M., Hinderks, A., & Thomaschewski, J. 2017. Design and evaluation of a short version of the user experience questionnaire (UEQ-S). International Journal of Interactive Multimedia and Artificial Intelligence, 4 (6), 103-108. https://doi.org/10.9781/ijimai.2017.09.001

Open-Source Virtual Reality for Vocational Training and Career Guidance

Mikhail Fominykh, and Ekaterina Prasolova-Førland

Norwegian University of Science and Technology, Trondheim, Norway

Corresponding author: mikhail.fominykh@ntnu.no

Keywords: Virtual Reality, Open-Source, Vocational Training, Career Guidance

Introduction

The software development project presented in this paper aims to address the challenge of motivating and empowering young job seekers in exploring workplaces, professions, and careers. The main target audience of the project are young job seekers who can be aided in selecting a career path at school or a welfare center, choosing the first or a new occupation, often after a period of being unemployed. Another target audience includes potential vocational education students who can be guided through workplaces and basic training in the transition from school to studying for blue-collar professions.

We earlier explored the potential of Virtual Reality (VR) for career guidance and counseling for young job seekers. Most of the participants in our study (79% of job seekers and 97% of welfare professionals) agreed or strongly agreed that VR apps should be made available to job seekers to explore workplaces and professions (Prasolova-Førland, Fominykh, and Ekelund 2019).

We consider VR to be a new and efficient way to communicate with our target audiences and to facilitate their engagement and a safe workplace exploration experience. VR can help the users to explore and train in unfamiliar situations in a safe setting, thus mastering the same real-world situation with the goal of mastering the pathway to study or work. VR can help to fill the communication gap between the industries and the young generation, for example, by providing a cost-effective and low-threshold alternative or supplement to internship placements using innovative technologies with gaming elements.

We earlier introduced an original concept of 'Immersive Job Taste' – interactive virtual reality demonstration of a workplace that aims to give a feeling of going through an average workday of a professional with elements of basic training (Fominykh and Prasolova-Førland 2019). Although the concept includes the experience of workplace tasks and situations, the objective is not to learn how to perform, but rather *to experience how it feels to perform* such tasks.

Development process

In this section, we describe the software development methodology adopted in the project, development roles we identified for contributors, and the current development team.

Development methodology

The software development process in the project is organized following the agile software development approach. The process, however, does not strictly follow the modern methodologies, such as Scrum, often adopted by companies. Our process and the practices are influenced by the experimental research-and-development nature of the project and the complex group of people involved, small size of the team and the involvement of students.

We use several adjusted Agile practices. Our sprints are one month long, adjusted to the small size of the team and their part-time commitment. We use issue tracking, grouping them into milestones and using Canban boards to track progress. We perform code reviews, assigned to a sprint master selected from experienced developers. We use wiki for documentation.

Our main code repository[5] contains core functionality that includes multiple prefabricated assets, released as packages, and a template scene with a neutral workplace environment and demonstrators of the prefabricated assets. Each of the other repositories contains a project dedicated to simulating a specific workplace. Releases of these repositories are VR applications.

Development roles

The development process is purposefully set up for an open-source project with clearly defined roles for different types of contributors. We aim to lower the barrier for various stakeholders to join the open-source development. The main contributor roles include (a) researchers and developers who could contribute to the development of the core functionality, using the project as a testbed case. The experimental functionality can be packaged, applied in any workplace-specific applications, and evaluated with target group users. The roles also include (b) university teachers who could task groups of students to develop new workplace-specific applications based on the existing core functionality. The role of a (c) subject matter expert is designed to streamline the design of content scenarios and collect information for them.

Development team

Our development team consists of employees and students at three European universities and subject-matter experts from a wider range of organizations. In each university, both employees and students are engaged. The students are engaged both as contracted employees and as part of their course projects and theses. The university employees, who are mainly experienced VR developers and researchers, are responsible for the core functionality, designing study projects, defining requirements, and technically supervising students.

Students engaged in the project typically receive a task to develop a new workplace application with one or multiple typical tasks. In this case, our subject matter experts advise them about the content. Some students are tasked with designing a specific asset for the core functionality.

[5] https://github.com/vr4vet/vr4vet

Software

The software we develop is split into core functionality and workplace-specific applications. Our goal with this setup is to standardize the VR applications and make the development more efficient. The software is developed using the Unity engine[6] and the VRI framework[7].

Core functionality

The core functionality is developed separately and includes everything that is needed or potentially needed in multiple workplace-specific applications. Following this approach, we aim to maximize the reusability of core assets and to reduce the maintenance effort. In addition, releasing the core functionality with an open-source license simplifies the process of developing new workplace-specific applications for both internal and external stakeholders.

The core of the software is a Unity project with template scene and a set of prefabricated assets. The template scene contains a neutral workplace virtual space with a player and example use of all prefabricated resources of the core. A new developer can explore the core functionality using the template scene and start new workplace-specific Unity projects by cloning it.

The prefabricated assets[8] are released as packages and imported the workplace-specific projects. First, the assets standardize user interface, user interaction, and navigation around the VR space. All menus, diegetic and spatial user interface elements, navigation, teleportation, and the use of buttons on the VR controllers are standardized. Second, the structure and presentation of subject content are also standardized. Most information the user receives about the workplace environment, workplace tasks, and necessary skills is presented not only via standard means but also follows the same structure. Third, we use some standardized game mechanics, such as feedback and scoring system, which are both still in active development.

Blue Sector app

Blue Sector is a VR app for exploring blue sector workplaces and professions[9]. It is one of our workplace-specific applications that is currently being developed and is already playable.

The workplaces in the Blue Sector app include aquaculture, fish processing, and seaport infrastructure. For example, the app contains a scene with a fjord with fish cages and a fish feeding control station (Figure 34). The user can learn about aquaculture, explore fish cages, and play a game that poses a realistic challenge of observing fish behavior via underwater cameras and adjusting the feeding intensity based on fish behavior.

[6] https://unity.com/
[7] https://wiki.beardedninjagames.com/
[8] Description of the prefabricated assets in VR4VET: https://github.com/vr4vet/vr4vet/wiki
[9] https://github.com/vr4vet/Blue-Sector

Figure 34. Simulated fish feeding control station in the Blue Sector app.

Acknowledgement

The work presented in this publication has been done withing the EU-funded project Virtual Reality for Vocational Education and Training - VR4VET[10], partially funded by the European Union's Erasmus Plus program, grant agreement 2021-1-NO01-KA220-VET-000028033. This publication reflects the views only of the authors, and the European Commission cannot be held responsible for any use which may be made of the information contained therein.

References

Fominykh, M. and Prasolova-Førland. E. 2019. Immersive Job Taste: a Concept of Demonstrating Workplaces with Virtual Reality. 4th Workshop on K-12+ Embodied Learning through Virtual & Augmented Reality (KELVAR), Osaka, Japan. IEEE. https://doi.org/10.1109/VR.2019.8798356.

Prasolova-Førland, E., Fominykh, M., & Ekelund, O. I. 2019. Empowering Young Job Seekers with Virtual Reality. IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, pp. 295-302, https://doi.org/10.1109/VR.2019.8798179.

[10] https://erasmus-plus.ec.europa.eu/projects/search/details/2021-1-NO01-KA220-VET-000028033

Measuring and Analyzing Differences of Operators’ Behavior Between Real and Immersive Workstations

Renjie Zhang, Jelena Petronijevic, Alain Etienne, and Jean-Yves Dantan

Arts Et Métiers Institute of Technology, Université de Lorraine, LCFC, HESAM University, Metz, France

Corresponding author: renjie.zhang@ensam.eu

Keywords: Operators Behavior, Human Factors, Production System Design, Virtual Reality, Data Analysis

Abstract

With the widespread use of Virtual Reality (VR) in various industries and the great results it has achieved, the integration of VR systems into production and manufacturing is an inexorable trend. However, there is no evidence that the operator’s behavior in the immersive environment corresponds to their behavior in real manufacturing situations. Therefore, in order to evaluate the effectiveness of VR, this paper implements a series of experiments to effectively measure the differences in operator behavior between immersive and real workstations. The main contribution of this paper is the evaluation and analysis of operator behavior in workstations with different environments. The results provide support and reference for future applications of virtual reality in the design of production systems.

Introduction

Advancements in computer technology have empowered Virtual Reality systems, making them invaluable in diverse domains, such as immersive gaming, telemedicine, and architectural planning (López Ríos et al., 2020). VR has proven its worth in product design, factory layout planning, and industrial optimization (Kose et al., 2019), becoming a fundamental economic tool (Matsas & Vosniakos, 2017). It has also demonstrated potential in interactive user behavior analysis for gaming. VR provides engineers with novel ways to visualize, interact, integrate, evaluate, and improve their designs (Havard et al., 2019). However, it's crucial to address biases of operators' behavior that may arise due to differences between real and virtual environments when integrating VR into production systems. Blind integration without reliability and effectiveness assessment could render the effort meaningless.

Based on these potential differences in operator behavior, researchers and scholars in many fields have been actively exploring the issue. For example, (Alejandro Huerta-Torruco et al., 2022) evaluated the effectiveness of using VR to train operators in a manufacturing system, where operator behavior was compared between two modes: laptop interface (LI) and virtual reality interface (VRI). In the medical domain,(Canessa et al., 2019). compared subjects' gait parameters in VR and real environments. The analysis of the considered gait parameters shows significant differences in the peak swing velocity, the step count and the cadence. This can be explained by the fact that subjects wore a cabled head-mounted display (HMD) for the virtual conditions, and this could be a possible factor that introduces differences in the walking behavior with respect to a real-world situation. Moreover, most of the subjects have never experienced walking in immersive VR before, thus it should be further analyzed whether these differences disappear through learning after a longer period of training.

It is evident from the field of related research that there are few studies in production and manufacturing to measure and evaluate operator behavior in real and immersive workstations. Therefore, the contribution of this paper is dedicated to filling the research gap in this area by performing experiments to evaluate the differences in operator behavior when performing assembly tasks between immersive and real workstations.

Implementation of the experiment

This paper designs and implements experiments to measure and evaluate differences in operator behavior between real and immersive workstations. In addition to the different workstation environments for the comparison experiment, we considered other factors that may affect operator behavior. Thus, the predictive model of operator behavior parameters should include two response functions and a residual error term. The first response function, $\tilde{f}$, describes the impact factor of the immersive or real workstation, while the second response function, $\tilde{g}$, represents the impact of other factors that may influence the operator behavior: the learning factor and the order of workstations. The residual error term, ε, represents any remaining uncertainties that cannot be accounted for by the two response functions.

$$\boldsymbol{Y_{operator\ behavior}} = \tilde{\boldsymbol{f}}(\boldsymbol{X_{immersive/real\ assessment}}) + \tilde{\boldsymbol{g}}(\boldsymbol{Z_{experimental\ factors}}) + \boldsymbol{\varepsilon}$$

	Constant	A	B	C	AB	AC	BC	ABC	Response
1	1	-1	-1	-1	1	1	1	-1	Y1
2	1	1	-1	-1	-1	-1	1	1	Y2
3	1	-1	1	-1	-1	1	-1	1	Y3
4	1	1	1	-1	1	-1	-1	-1	Y4
5	1	-1	-1	1	1	-1	-1	1	Y5
6	1	1	-1	1	-1	1	-1	-1	Y6
7	1	-1	1	1	-1	-1	1	-1	Y7
8	1	1	1	1	1	1	1	1	Y8

A: without learning -1
with learning 1
B: Immersive -1
Real 1
C: Immersive first -1
Real first 1

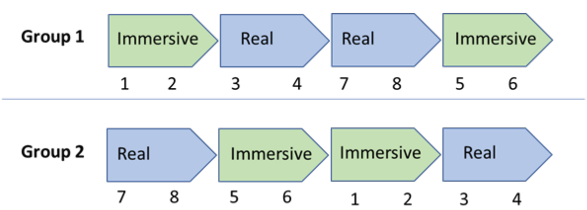


Figure 35. The process of designing the comparison experiment with Hadamard matrix.

In this paper we use the Hadamard Matrix as a mathematical method to plot the experimental process. A Hadamard matrix is an n*n matrix with elements ± 1 and mutually orthogonal columns. As shown in the table of Figure 35. Letters A, B, C represent the four factors that may influence the difference in operator behavior: operating experience, operating environment, and the order of workstation, correspondingly. By arranging and combining these three factors, 8 possible responses can be obtained. Then, the outcome of these 8 responses can be used to design the experimental process as shown in the bottom diagram of Figure 35.

Figure 36. Immersive workstation and Real workstation

The comparison experiment consisted of 9 participants, equally divided into two groups, with the first group starting at the immersive workstation and the second group starting at the real workstation. Each group performed 4 scenarios. In each scenario, the task of assembling 10 products was performed twice. Regarding the experimental equipment used in this experiment, Meta Oculus Rift S is selected as the VR hardware equipped with SkyReal v1.15 software to provide the operating environment of an immersive workstation. The participants' experimental scenarios in the immersive and real workstations are shown in Figure 36.

Data processing and analyzing results

In terms of data processing, this paper uses the duration as the response parameter for the analysis. The results of the analysis of assembly duration with the help of Hadamard matrix are shown in Figure 37. In addition, the professions of the participants are grouped in this paper in order to better analyze the differences of the results between groups. Time units used for data results are in seconds.

Profession	Duration	A	B	C	AB	AC	BC	Group	ABC
Technical staff	32,8625	-0,8875	0,0625	-1,7625	-0,1375	0,6375	-2,0125	Real	0,4375
Technical staff	28,0625	-2,0375	0,2875	-2,0625	-0,5125	1,1375	-0,2875	Real	1,1125
Technical staff	39,9053	-0,5553	-0,2446	-1,9196	-0,5553	-0,8803	4,0304	VR	-0,3803
Technical staff	34,0375	-2,0375	-0,0625	-1,3125	-0,9875	-0,3875	4,1875	VR	-1,4375
Young Engineer	42,1875	-0,8625	0,2125	-5,4375	0,6125	-1,1375	4,8375	Real	1,8875
Young Engineer	46,3625	-2,7625	-0,7625	-5,7625	**-2,4375**	1,6125	-4,4375	VR	-0,6125
Young Engineer	38,2	**0,6**	1,925	-1,2	0,725	1,15	-1,975	VR	0,575
Teacher	30,7375	-1,1125	1,5875	-0,9125	**-1,7625**	0,8875	5,8375	Real	0,0375
Teacher	35,2	**1,125**	0,15	-0,675	1,075	1,15	-3,775	VR	0,75

Figure 37. The values of each impact factor which calculated by Hadamard's matrix

In terms of participants' professions, the assembly duration of technical staff and teachers have mostly shorter assembly duration than young engineer. For Learning Factor, A, the majority of the participants' assembly time decreases due to the learning factor. It makes sense that the assembly duration with learning would be shorter than the one without learning. However, there are two abnormal values in the list, 0.6 and 1.125. This is most likely due to the hot weather on the day of the experiment, which caused fatigue in the experimenters and led to an increase in the assembly time. However, factor C - workstation order has a significantly impact on duration, with assembly duration decreasing for each participant regardless of which workstation they started the experiment from. This may be due to the influence of the learning factor, as the number of products assembled increases, the participants' assembly duration decreases.

Regarding the effect of factor, A+B (learning factor + workstation environment) on assembly duration, there are two noticeable abnormal values of -2.4375 and -1.7625. The main reason for the significant difference between these two values and the other values may be that the participants are more capable of learning, so that the assembly duration is significantly reduced in the later stage of the experiment compared to the beginning stage of the experiment. Among the different profession groups, the effect of the combined factor B+C (workstation environment + workstation order) on assembly duration displays a significant correlation with workstation order. However, this property is not significant for either factor B or C alone.

Conclusion

This paper bridges the gap in research in the production industry by experimentally measuring and analyzing the differences in operator behavior between real and immersive workstations. Based on the scientific issues, the paper describes the comparison experiment, and explains the methods of the experimental implementation. The data collected are then processed and analyzed. It is concluded that there are partial differences in operator behavior between immersive and real workstations, which are due to a number of factors. When using virtual reality to assist in the design of production system, factors that can make a difference to operator behavior should be avoided as much as possible. In addition to the duration analyzed in this paper, we also recorded the movement data and subjective feelings of the participants in our experiments. In the future work, we will conduct a more comprehensive analysis of operator behavior based on these data.

References

Canessa, A., Casu, P., Solari, F., & Chessa, M. 2019. Comparing real walking in immersive virtual reality and in physical world using gait analysis. Proc. 14th Int. Joint Conf. Comp. Vision, Imaging and Computer Graphics Theory and Applications - VISIGRAPP, 2. https://doi.org/ 10.5220/0007380901210128

Havard, V., Jeanne, B., Lacomblez, M., & Baudry, D. 2019. Digital twin and virtual reality: a co-simulation environment for design and assessment of industrial workstations. Production and Manufacturing Research, 7(1). https://doi.org/10.1080/21693277.2019.1660283

Huerta-Torruco, V. A., Hernandez-Uribe, O., Cardenas-Robledo, L. A., & Rodriguez-Olivares, N. A. (2022). Effectiveness of virtual reality in discrete event simulation models for manufacturing systems. Computers & Industrial Engineering, 168, 108079. https://doi.org/10.1016/j.cie.2022.108079Get rights and content

Kose, A., Tepljakov, A., Abel, M., & Petlenkov, E. 2019. Towards assessment of behavioral patterns in a virtual reality environment. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11613 LNCS. https://doi.org/10.1007/978-3-030-25965-5_18

López Ríos, O., Lechuga López, L. J., & Lechuga López, G. 2020. A comprehensive statistical assessment framework to measure the impact of immersive environments on skills of higher education students: a case study. Int J Interact Des Manuf 14, 1395–1410. https://doi.org/10.1007/s12008-020-00698-1

Matsas, E., & Vosniakos, G. C. 2017. Design of a virtual reality training system for human–robot collaboration in manufacturing tasks. Int. J. Interact. Des. Manuf. 11, 139-153. https://doi.org/10.1007/s12008-015-0259-2

XR Standardization – a Global Overview

Christoph Paul Runde

EuroXR Association / Virtual Dimension Center (VDC)

Corresponding author: christoph.runde@vdc-fellbach.de

Keywords: virtual reality, augmented reality, mixed reality, extended reality, norming, standardization, guidelines, recommendations, norms, specs, specification, industrialization

Scope

The landscape of XR norms, standards, guidelines and recommendations (hereafter summarized for the sake of simplicity as: XR norms) is extremely broad, scattered and confusing today. It was already evident from the preliminary work that dozens of organizations are working on the topic, some of which in turn host several working groups. At least hundreds of documents are estimated to be relevant. Today, there is no institution that comprehensively sifts, classifies and transfers the knowledge on XR standards available to the public. This work attempts to create transparency in the field of XR standardization.

Due to the scattered nature of this topic, the research was conducted with different starting points. Standards organizations, but also associations of XR user industries were identified and searched for XR committees, XR working groups and XR standardization projects. Based on further known XR standards, it was again possible to infer working groups, standards organizations and associations. The search for documents was also initiated via (university) libraries and a free WWW search. This process has been run through cyclically many times. As a result, completed lists of standards organizations, working groups, ongoing activities and XR standards are available.

Norming and Standardization

Standardization is generally the unification of products, components or processes to one or a few variants. It is a standardization of objects according to certain templates. The procedures for standardization are norming and typification. The term can be applied to various fields. In the economic fields: Manufacturing area norming and typification of parts, intermediate or final products.

Standardization refers to the formulation, issuing and application of rules, guidelines or characteristics by a recognized organization and its standards bodies. They shall be based on the established results of science, technology and experience and aim at promoting optimal benefits for society. The specifications shall be established by consensus and adopted by a recognized institution. Recognized standardization institutions are, for example, ISO, IEC, EN, DIN.

A. Arguments in favor of Standardization

The main purpose of standardization is to save costs and simplify work. Standardization leads to an increase in market transparency and a reduction in costs (in manufacturing costs, information costs, transaction costs, shipping costs, distribution costs, switching costs). By using standards, manufacturers can concentrate on the really innovative aspects of their products.

On the part of the buyers of standardized products, services and services, the reduced switching costs and compatibility are particularly interesting: purchased products and services become much easier to exchange and technically compatible, and thus integrable. This intensifies competition, which must lead to price degression and increased performance.

A very important advantage of standardization is the attainment of a certain legal certainty for products brought onto the market. Of particular legal interest is the so-called reversal of the burden of proof: in the event of damage to a non-standardized product, the manufacturer must prove that the product was developed without defects. If, however, a case of damage occurs with a standardized product, the manufacturer is deemed to have acted in accordance with the state of the art. In this case, the customer must prove that the manufacturer acted incorrectly. Here, the burden of proof is reversed.

B. Arguments against Standardization

However, the positive aspects of standardization also have possible disadvantages.

For the success of a product, a service and thus the entire company, a USP ("unique selling proposition") should be given. Thus, it may seem problematic that a standard solution can be the basis of a good USP, as it has to stand out from the crowd. This inevitably means that the special added value of the product, instead of coming from a standardized value chain, can only come from a deviation from the standard.

If one analyses existing standards, it becomes clear that the criteria formulated are always minimum requirements (i.e. what the customer should be able to expect anyway).

The consistent pursuit of standards can thus lead to companies always finding themselves at the bottom of the range and even giving up their unique selling propositions. Blue Ocean strategies include - somewhat abbreviated - in particular the omission of learned features of a product or service if this makes it possible to create significant added value for the customer elsewhere. Here it makes sense not to meet standards if this serves the unique selling proposition and the success of the product.

Standardization creates the dangers of schematization and loss of flexibility. It can happen that aspects that cannot be standardized are neglected. Standardization is a coordination-intensive process that incurs high costs and normally takes several years to complete. This results in a framework that is too static. Standardization can lead to a restriction of creative freedom. This stifling of innovation fields is undesirable, especially in the environment of the development of innovative technologies.

Approach of this Work

The landscape of XR norms, standards, guidelines and recommendations (hereafter summarized for simplicity as: XR norms) is extremely broad, scattered and confusing today. It was already evident from the preliminary work that dozens of organizations are working on the topic, some of which in turn host several working groups. At least hundreds of documents are estimated to be relevant.

Today, there is no institution that comprehensively sifts, classifies and transfers the knowledge on XR standards available to the public. This work shall achieve this task.

A. Standardization in the V/AR context

Virtual Reality (VR) is a spatial user interface for 3D data. VR can be defined as a computer-generated, real-time 3D environment in which one or more persons are immersed by uniquely locating them in the spatial coordinate system of the 3D scene via position detection. Only in this way does the perception of the 3D content (change of perspective, direction-dependent hearing, scanning, etc.) react as we are used to from natural reality. This enables people to better grasp the 3D content presented and understand it in its spatiality. With the help of VR, 3D environments can thus be perceived better than with simple desktop computer systems. At the same time, VR supports spatial interaction with the presented 3D data. The basic prerequisite for the meaningful use of VR is spatially-geometrically complex 3D data. Augmented Reality (AR) is the superimposition of the natural perspective of sight with (3D) computer graphics. AR thus merges a virtual environment with reality. This can be useful for assistance systems or target/actual comparisons (digital plan versus physical reality). Mixed Reality (MR) is the simultaneous presentation of natural and artificial sensory stimuli, mostly digital visualization in combination with physical-haptic interfaces.

Virtual reality and augmented reality are not new methods: the first implementations of VR began in the 1960s at the latest, those of AR in the 1970s at the latest. VR and AR (also V/AR or XR) are cross-cutting technologies and methods that can encompass a huge number of knowledge domains. These include perception/cognitive psychology, work sciences, computer graphics, acoustics, haptics, user interface design, hardware development, software development, etc. In addition, for practical use, there is often concentrated knowledge from the application field (such as design, maintenance, ergonomics, industrial engineering, marketing communication, etc.) and from the industry. This diversity of subject areas is also directly reflected in the consideration of the standardization fields.

B. Monitoring organizations, XR standards and standardization projects

The identification of all relevant XR standardization organizations, XR standards and standardization projects implies very extensive research tasks. We did so by active, personal exchange with representatives of standardization organizations and by participation in relevant events at which XR standardization.

Due to the scattered nature of this topic, the research was conducted with different starting points. Standards organizations, but also associations of XR user industries were identified and searched for XR committees, XR working groups and XR standardization projects. Based on further known XR standards, it was again possible to infer working groups, standards organizations and associations. The search for documents was also initiated via (university) libraries and a free WWW search. This process has been run through cyclically many times. As a result, completed lists of standards organizations, working groups, ongoing activities and XR standards are available.

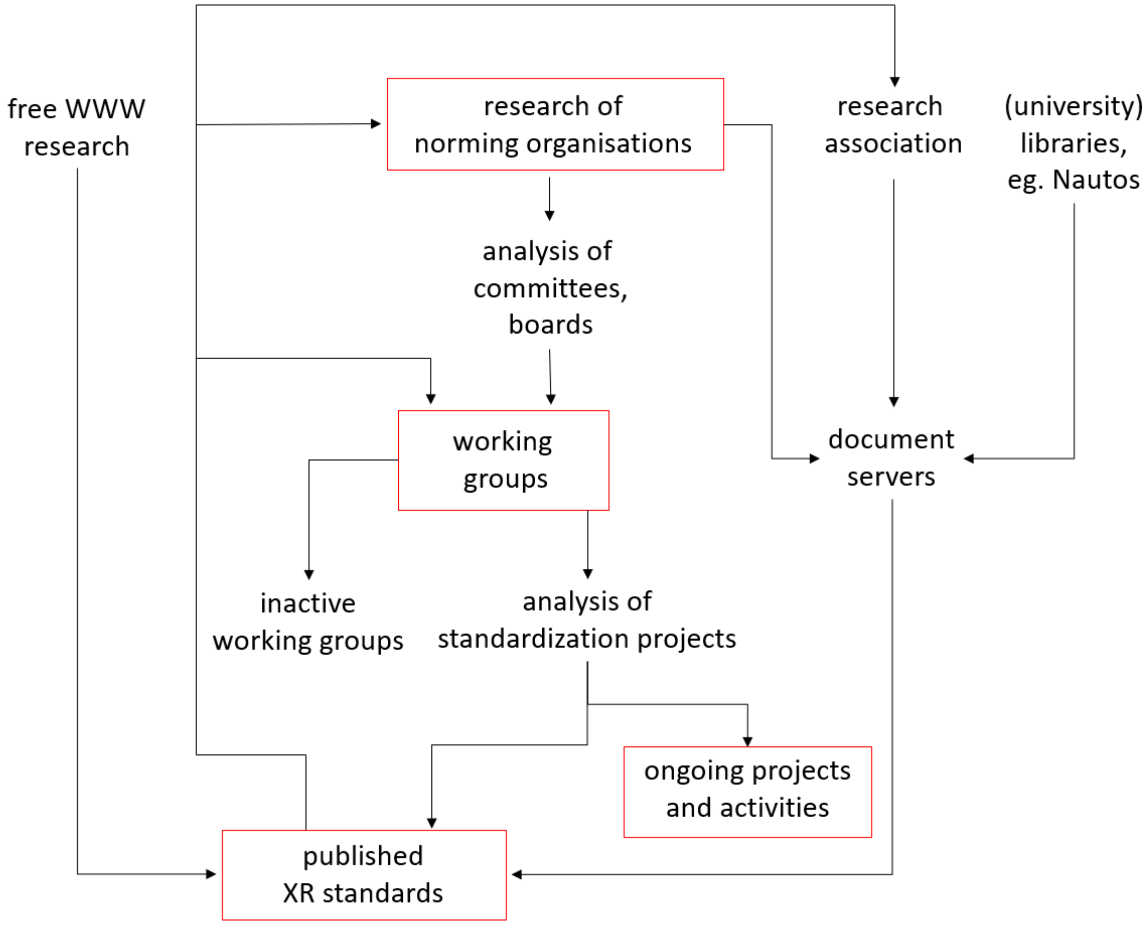


Figure 38. Research strategy for XR norms and XR standards

Thematic classification of norms

A. Categorization of XR standards

From the knowledge of the topics dealt with in the XR standards and working groups, categories are formed after content analysis that are sufficiently detailed (and thus enable selectivity), but at the same time still manageable in number. We generated 7 categories and 31 sub-categories (c.f. Figure 39).

Fundamentals of VR AR	▪ basics / terms ▪ XR management ▪ regulation / ethics / governance ▪ evaluation / conformity
XR management	▪ basics / terms ▪ XR management ▪ regulation / ethics / governance ▪ evaluation / conformity
UX, ergonomics, human factors	▪ user experience (UX) / user interface design (UID) ▪ ergonomics / usability ▪interaction / pattern
Coding, mapping, interoperability, communication	▪ API application programming interface ▪ formats / coding / compression ▪ communication / interoperability

	▪ mapping
Graphics software, CGI	▪ graphics software / algorithmics
Hardware: optics, haptics, acoustics, tracking, mobile	▪ graphics hardware / optics / displays ▪ haptics ▪ auralisation / audio / acoustics ▪ tracking / sensor technology / motion capturing ▪ mobile XR ▪ other XR hardware
XR applications	▪ human models ▪ content creation / modelling ▪ applications: education / training ▪ applications: design / engineering ▪ applications: health ▪ applications: assistance ▪ applications: collaboration ▪ applications: marketing ▪ applications: entertainment / culture ▪ applications: geo & construction ▪ applications: other

Figure 39. Categories for XR norms and XR standards

The sub categories are further explained under sections b) to h).

When analyzing what areas of XR have already been subject to standardization, we find that many documents have been edited on formats/coding, graphics software, however much less on certain application fields of on interaction (c.f. Figure 40). Figure 41 shows a grouped perspective on this issue with according numbers.

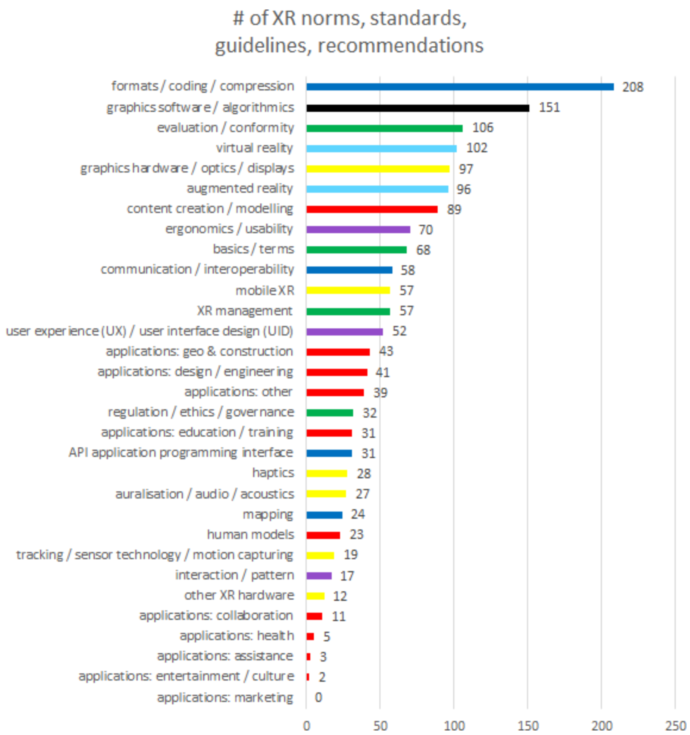


Figure 40. Number of existing XR norms and standards in relation to their topic

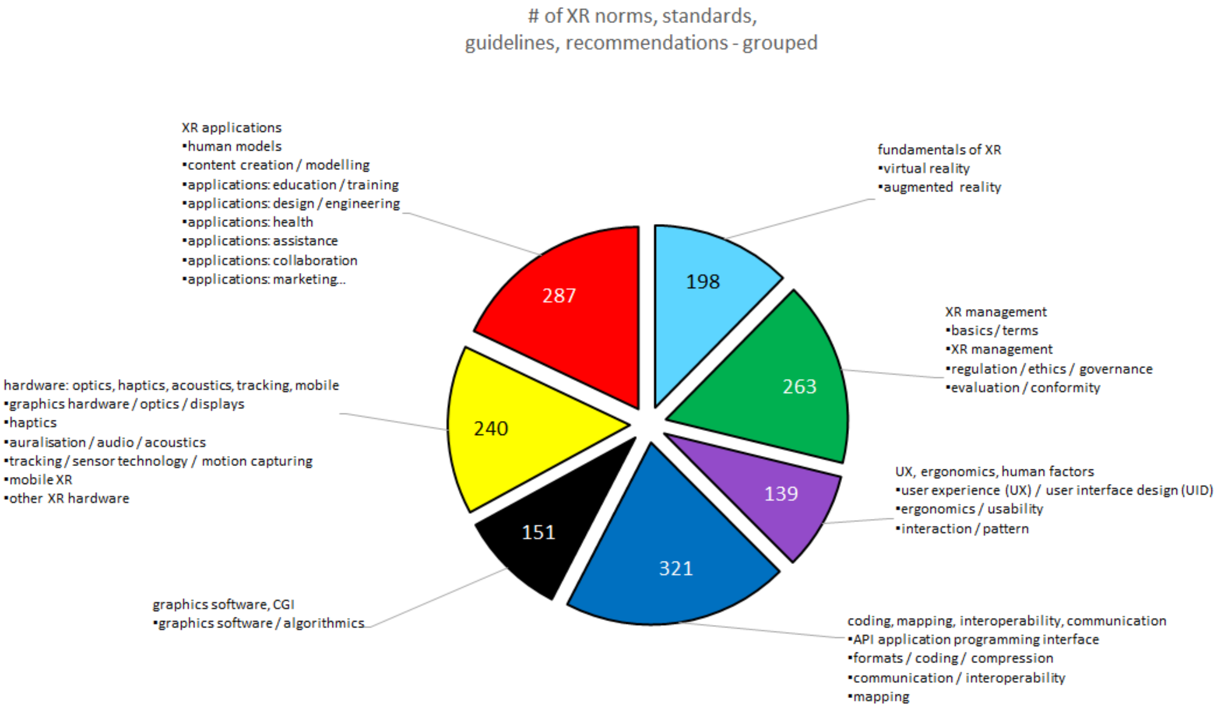


Figure 41. *Number of XR norms, standards, guidelines, recommendations - grouped*

B. Area 1: Fundamentals of XR

In the first topic area of "*Fundamentals of XR*", terms, definitions, characteristics, taxonomies are discussed. The aim here is to use uniform terminology and definitions: only if a common language is spoken can meaningful action be taken together. Relevant standards include e.g.:

- CTA: Definitions and Characteristics of Augmented and Virtual Reality Technologies
- IEEE P2048.1 Standard for Virtual Reality and Augmented Reality: Device Taxonomy and Definitions
- IEC 63203-101 Wearable electronic devices and technologies – Part 101-1: Terminology

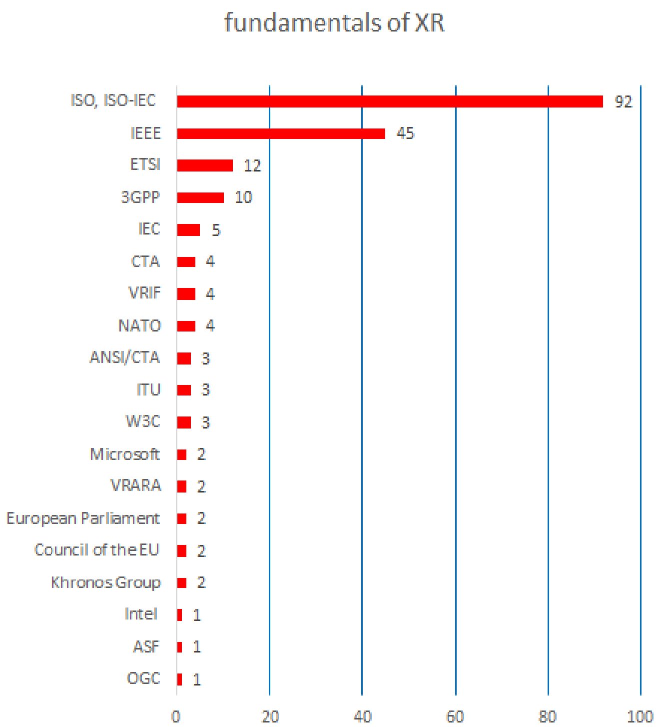


Figure 42. *Number of standards publications on fundamentals of XR by organization*

C. Area 2: XR management

In the V/AR management subject area, organizational, administrative and managerial aspects are central. Relevant standards include e.g.:

- ANSI/CTA: Recommendations and Best Practices for Connection and Use of Accessories for XR Technologies
- Council of the European Union: Metaverse - Virtual World, Real Challenges
- ETSI: Augmented Reality Framework (ARF). Industrial use cases for AR applications and services
- ISO-IEC. AR/VR safety-- guidance on safe immersion, set up and usage

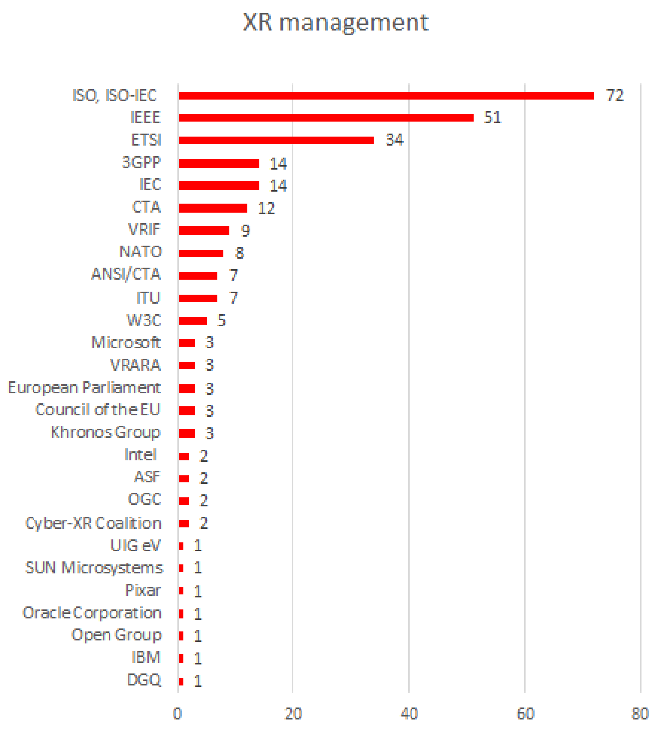


Figure 43. *Number of standards publications on XR management by organization*

D. Area 3: UX, ergonomics, human factors

Compared to standard desktop interaction systems, consisting of mouse, keyboard, mouse pointer and window display, V/AR involves much more complex interaction mechanisms, which makes separate standards necessary. However, there is still no generally accepted or even adopted, binding set of rules. For the most part, only guidelines and practicable examples (best practices) are available. Examples are:

- ISO-IEC: Information technology for learning, education, and training – Human factor guidelines for virtual reality content
- IEEE: Standard for Head-Mounted Display (HMD)-Based Virtual Reality(VR) Sickness Reduction Technology
- ISO: Ergonomics of human-system
- ITU: Influencing factors on quality of experience for virtual reality services

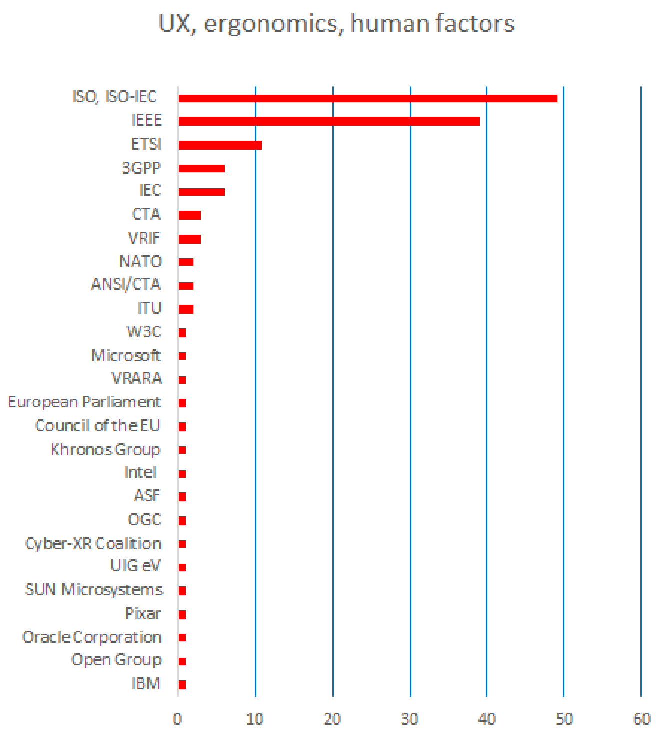


Figure 44. *Number of standards publications on UX, ergonomics, human factors by organization*

E. Area 4: coding, mapping, interoperability, communication

The topic of interoperability and communication deals with aspects that have the background of allowing V/AR components to work together with other systems, be it legacy IT or physical reality components. Some examples of relevant standards are:

- 3GPP Virtual Reality (VR) streaming interoperability and characterization
- ETSI Augmented Reality Framework (ARF) Interoperability Requirements for AR components, systems and services; Part 1: Overview
- IEEE Standard for VR and AR: Interoperability between Virtual Objects and the Real World
- ITU Interoperability testing requirements for a virtual broadband network gateway

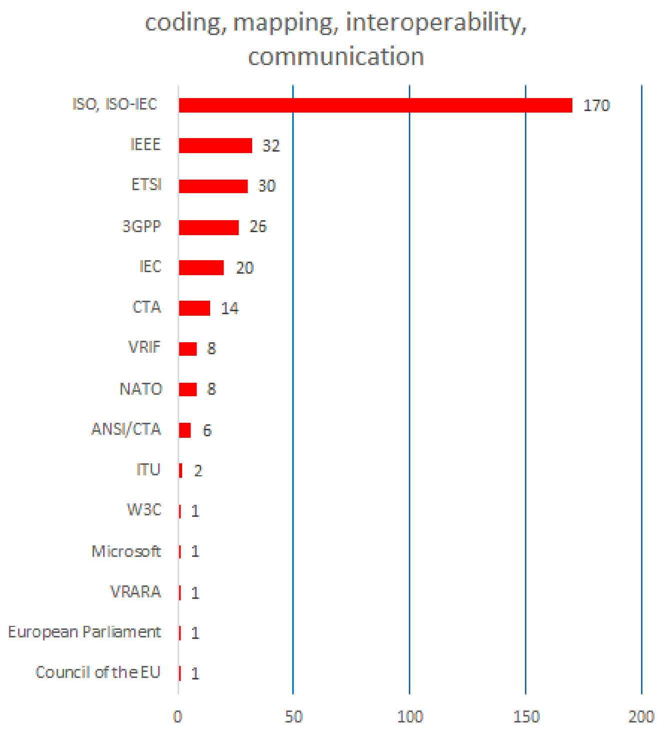


Figure 45. *Number of standards publications on coding, mapping, interoperability, communication by organization*

F. Area 5: graphics software, CGI

There is a whole range of different mathematical description methods for 3D geometries, which differ in terms of various properties (generality, precision, one-to-one uniqueness, speed of representation, etc.). Accordingly, there are numerous different 3D data formats. These certainly move on different levels of abstraction (up to semantic aspects) and thus cross over into a border area to the subject area of content and applications. The relevant standards include, for example:

- 3GPP: VR profiles for streaming applic.
- VRIF: Volumetric Video Guidelines
- ISO-IEC: Information technology - Computer graphics, image processing and environment data representation
- ISO-IEC: Extensible 3D (X3D)

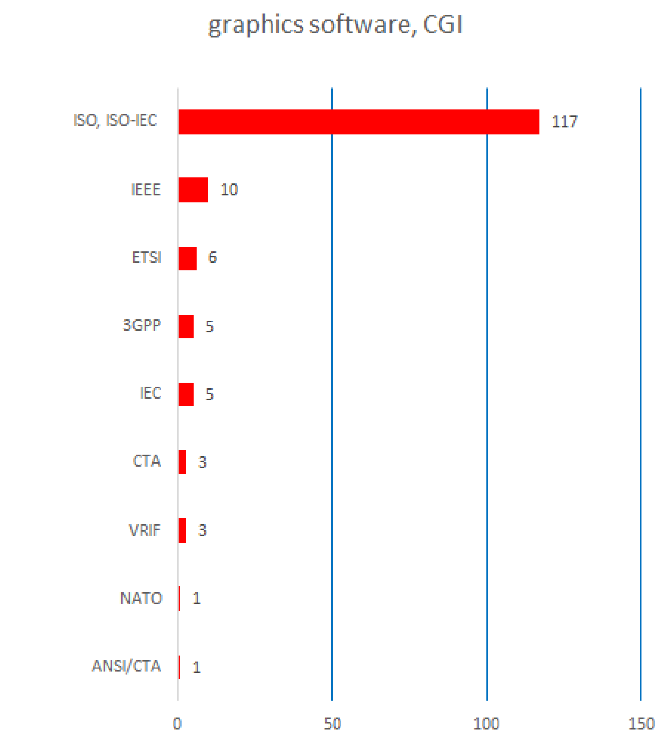


Figure 46. *Number of standards publications on graphics software, CGI by organization*

G. Area 6: hardware: optics, haptics, acoustics, tracking, mobile

Hardware standards deal with optical properties of graphic output systems, haptic properties of tactile, haptic and kinaesthetic output systems, auditory properties of acoustic output systems and measurement methods for recording them. Other hardware areas not explicitly mentioned may also be relevant to the V/AR environment. Examples are:

- IEC: Specific measurement methods for AR type - Image quality
- IEC: Touch and interactive displays - Measuring methods of touch displays
- Khronos Group: OpenSL ES standard for three-dimensional audio systems
- IEC: Wearable electronic devices and technologies

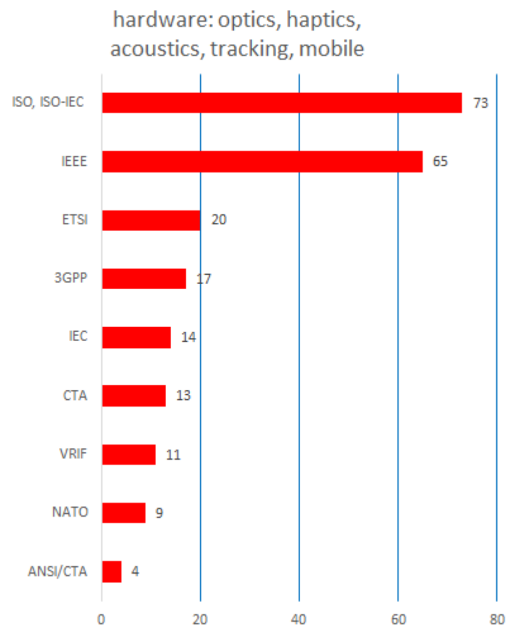


Figure 47. Number of standards publications on hardware: optics, haptics, acoustics, tracking, mobile by organization

H. Area 7: XR applications

The norms and standards in the context of content and applications go beyond the pure mathematical descriptions of geometries, scene graphs, colors, materials, etc. They define the application of 3D scenes in a specific context and must therefore also provide meaning-related information. They define the application of the 3D scenes in a specific context and must therefore also provide meaning-related information. Some important applications of V/AR today are product development, industrial engineering or training. Examples are:

- ETSI: Augmented Reality Framework (ARF). Industrial use cases for AR applications and services
- NATO: Guidance in the use of simulation and virtual prototyping in ship design
- ISO: Clothing - Digital fittings - Attributes of virtual garments
- ISO: Health informatics - Reference model for VR based clinical practice simulation
- XRA: Designing Immersive Learning for Secondary Education

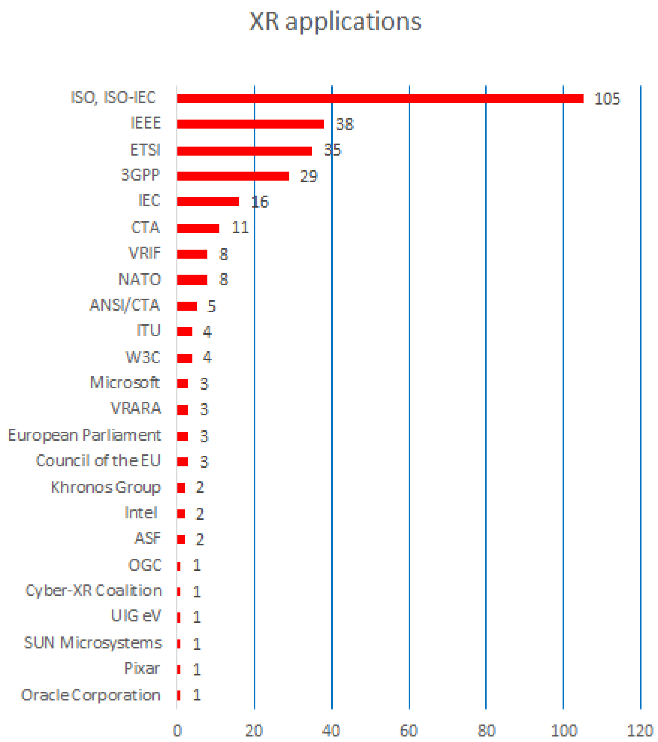


Figure 48. Number of standards publications on XR applications by organization

XR Standardization Organizations

A. Relevant Stakeholder Organizations

We found more than 40 organizations that are actually active in V/AR standardization and that already published relevant documents. Figure 49 shows the most relevant organizations.

Figure 49. *XR standardization organizations today*

These above illustrated organizations published more than 600 relevant XR norms, standards, guidelines and recommendations. Figure 50 shows how many documents have already been released by them.

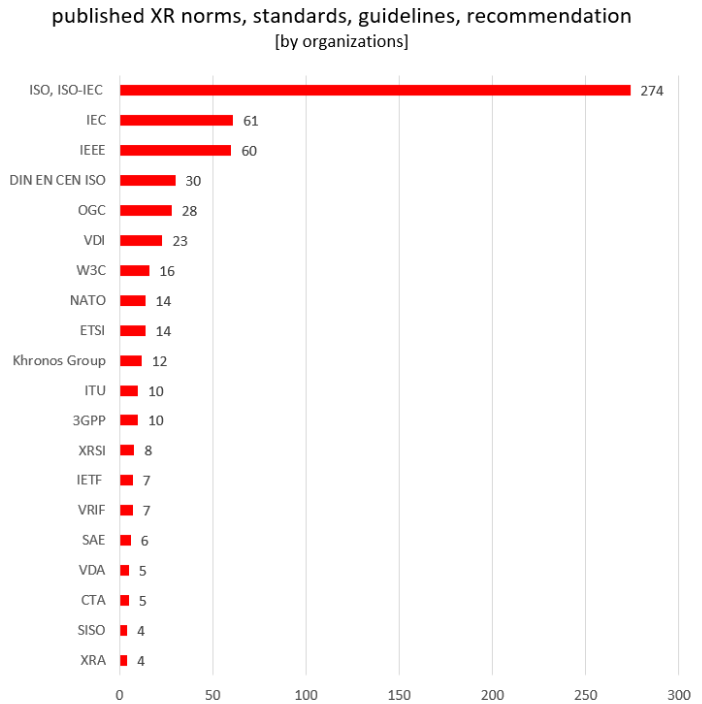


Figure 50. *Published XR norms, standards, guidelines, recommendation [by organizations]*

B. Focal areas of the work of the standards organizations

When analyzing the standards organizations' areas of activities, it is obvious that they focus on different priorities. Whereas e.g. ETSI, ITU, 3GPP are very active in the field of communication, IEC is intensively working on hardware, and NATO, SAE and VDI care for XR applications. We visualized those findings in the following spider net / radar diagrams.

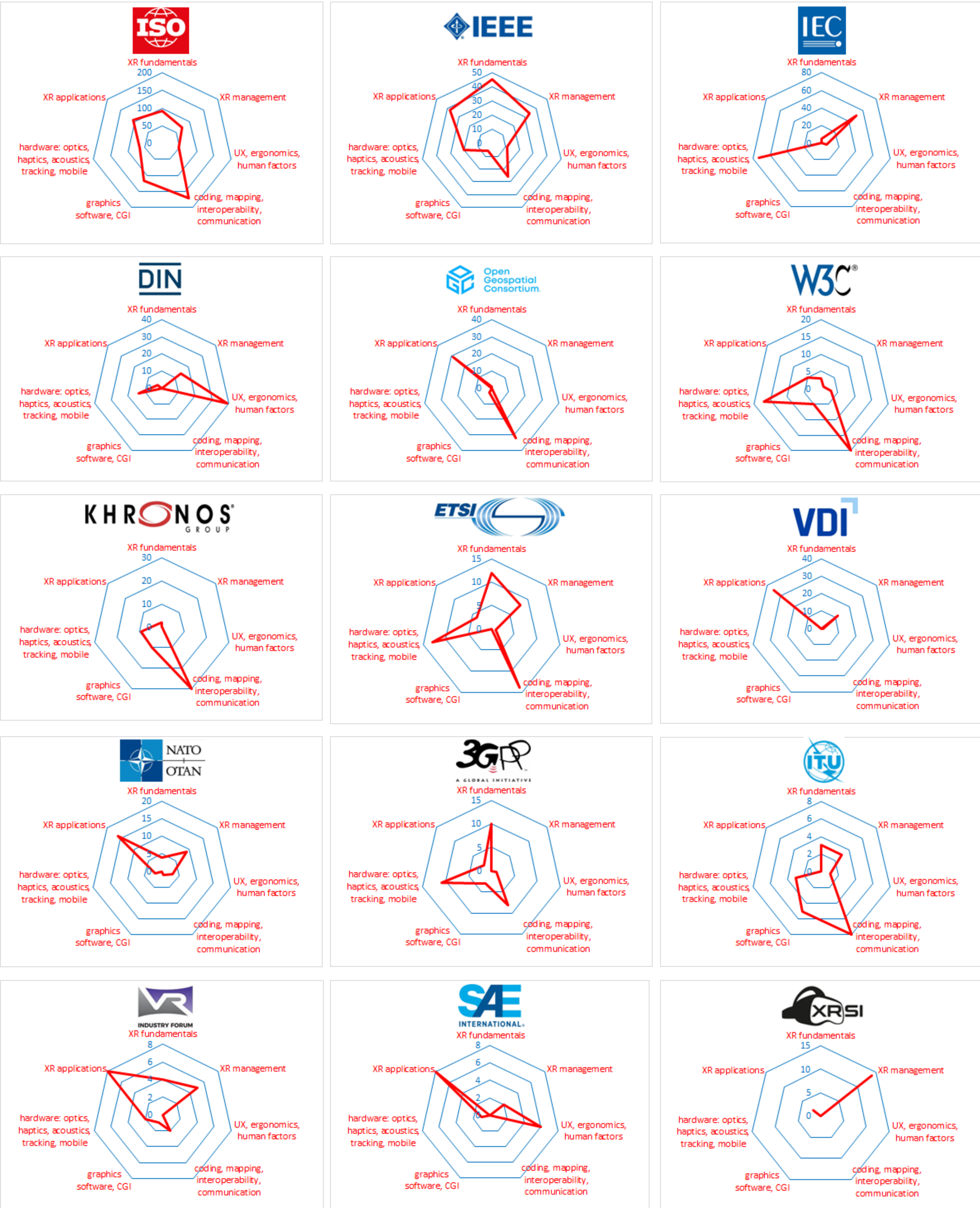


Figure 51. Focal areas of the work of the standards organizations

The Global XR Standards Monitor

We have created a web-based standards directory. This is based on the commercial library software ZOTERO. The directory now contains more than 640 entries on XR standards, guidelines and recommendations. The title of the document, the date of publication, the author and a short summary are stored in each case. The entry is also classified according to the categories presented above. The standards documents themselves are, of course, not part of the ZOTERO database, as they are protected by copyright.

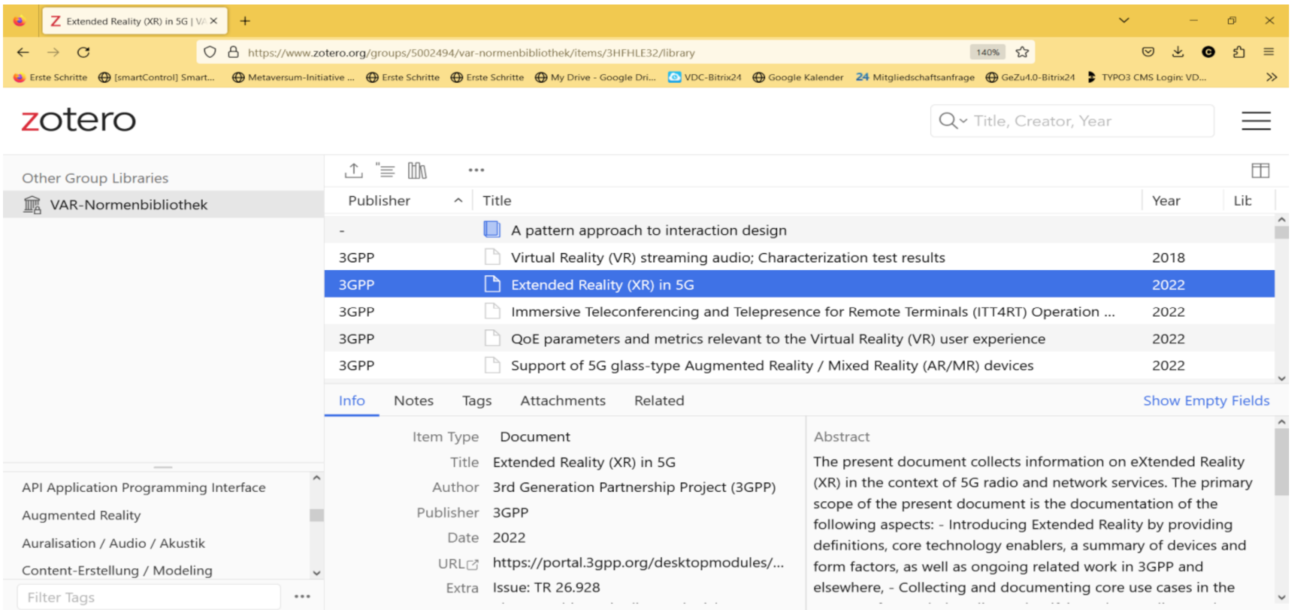


Figure 52. Web-based XR Standards Directory

Secondly, the database also contains a list of the currently active V/AR working groups in standardization. Users of the database can thus get a very good picture of which organisation is currently working on which XR topic. In this way, it should be much easier for interested parties to get started with the right topic. This section currently contains over 85 entries.

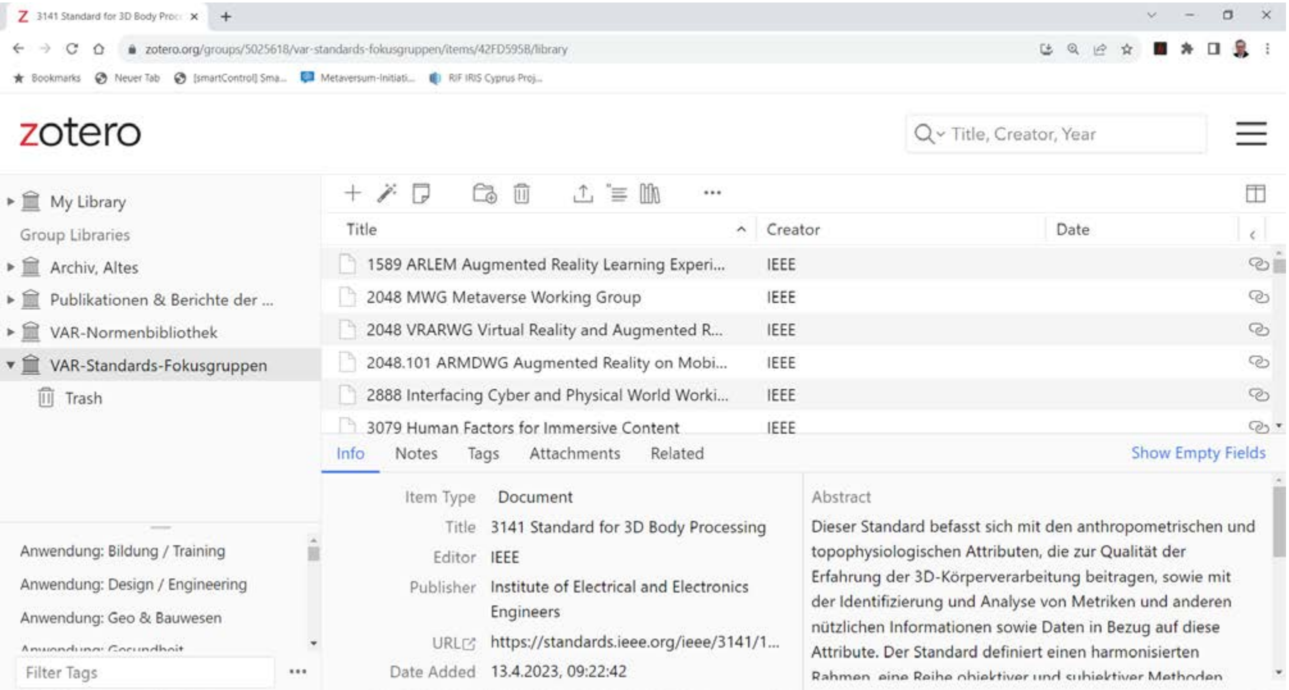


Figure 53. Active V/AR focus groups working actually on XR Standardization

VII. Further Demands for XR Standardization

The British Standards Institution (bsi) [bsi] published the paper "The Requirement for Standards in the VR and AR Sectors" in March 2018. In addition to an assessment of the current market development and the positioning of the British industry in the V/AR segment (focus: Creative Industries, Health, Safety, Skilled Workforce), some topics would also be mentioned in which the bsi would like to see more standardization activities. Overall, the bsi's focus is very much on the creative industries and their V/AR applications. Timmerer [Timmerer] gives an overview of standardization activities in the field of immersive media in "Immersive Media Delivery: Overview of Ongoing Standardization Activities". At the same time, he mentions some points where he sees a need for further standardization in the V/AR context. Stockhammer [Stockhammer] presents the MPEG MP20 Standardization Roadmap. Won Lee [Won Lee] created with "White Paper. Guidelines for Developing VR and AR Based Education and Training Systems", a guide for the development of V/AR-based training systems. In it, he also makes some recommendations on how to proceed. In his opinion, his approach is transferable to other fields of application of V/AR. Price [Price], in her paper "The role of international standards in virtual education and training systems", expands the view especially into neighboring subject areas when she deals with the topic of V/AR-based training. In particular, she advocates using the existing competences and standards there and adapting them to the V/AR application. Wajahat [Wajahat] explains in his lecture "New Proposal: Mixed methods User Experience Evaluation in AR/VR. A lean process for selecting appropriate UX evaluation methods and techniques in AR/VR" the proposal for the new ISO/IEC JTC 1/SC 24. In his presentation "Medical 3D Printing Scanning and Standards Requirements", Shim [Shim] deals with application-oriented standardization and standardized process chains in the work process of medical 3D image processing and 3D printing. In the "Augmentend5G" project [04], the project partners Aixemtec GmbH and oculavis GmbH from Aachen, together with Hella GmbH und Co. KG from Lippstadt, are developing new augmented reality applications for the production and assembly of optical systems under the leadership of the Fraunhofer Institute for Production Technology IPT. For the development work, the production machines and systems are integrated into a "Remote Expert Platform". The project team also wants to test the use of high-performance 5G mobile radio technology for data transmission in remote service. In his article "Standardization Could Be A Major Problem For Virtual Reality", Mirt [Mirt] raises awareness of the disadvantages of closed V/AR ecosystems, such as those promoted by the companies Facebook (Oculus Store), Sony (Playstation VR) and HTC (Steam platform).

Overall, the demands of the above-mentioned authors can be classified into the following five topics:

A. System integration

Mirt [Mirt] pleads for open V/AR standards such as WebVR at the time, today WebXR, in order to prevent the formation of closed technical ecosystems. Won Lee [Won Lee] also sees a need for standardization in the development of a system integration methodology for V/AR-based education and training systems as well as for V/AR-based health information systems and wearable systems.

B. Data transmission, interoperability

The bsi [bsi] calls for industry-wide standards for metadata and subtitles in V/AR content. These should be supported by VR platforms, content acquisition, software companies, content rights holders. Stockhammer [Stockhammer] and Timmerer [Timmerer] call for more efficient mapping (Projection) formats instead of today's "equirectangular" method for 360° content, as well as better encoding and encapsulation mechanisms for adaptive delivery of multimedia content. Price [11] advocates the use of ISO/IEC SC 24 (see above, graphical data processing), as well as SC 29 (data transmission, encoding / decoding) in the V/AR context. In the "Augmentend5G" project [04], the partners want to set the first industry standards for data conversion for AR.

C. Content

The bsi [bsi] calls for standards for content (for health and safety). The bsi continues to see advantages in standards for the placement of subtitles in VR content. Won Lee [Won Lee] identifies new areas of work in a framework for V/AR-based education and training systems. Price [11] advocates the use of ISO/IEC SC 36 (learning/education/training LET, along with associated data processing). LET information should be stored in a standardized format in a knowledge database. Shim [Shim] proposes a secure data processing process while respecting already existing legacy systems and standards (CT, MRI, DICOM, segmentation, 3D handling, STL imager, CAD/CAM, slicer/G-coder, 3D printer).

D. Guidelines for application

The bsi [bsi] calls for Best Practice Guidelines for the safe use of V/AR. Furthermore, standards and best practice guidelines for 360° video content were mentioned as necessary. In the "Augmentend5G" project [04], the partners would like to set the first industry standards for the software-supported creation of AR instructions.

E. Terms, quality, user experience (UX)

The bsi [bsi] sees a need for clarification of terms, nomenclature and quality assurance through the assignment of protected terms "VR", "AR", etc. (by the bsi). Timmerer [Timmerer] also sees a strong need for a standardized definition and treatment of aspects of user experience (UX) and quality of experience (QoE). Price [11] demands that suitable user interfaces must ensure the handling and control of the interaction with LET (see above) in virtual environments. Wajahat [Wajahat] advocates standardizing a UX evaluation of V/AR systems using several combined UX analysis methods.

Conclusion

A. Expectations to Policy Makers

Secluded technical V/AR ecosystems lead to higher customer loyalty due to the "customer lock-in effect". This hinders competition and reduces the pressure to innovate. Open standards such as WebXR (W3C, Mozilla) and OpenXR (Khronos Group) provide a remedy. The same applies to the lack of 3D data processing standards: they lead to increased workloads, may require customized solutions and thus also have a lock-in effect. The added value to be achieved with V/AR is reduced. Poor V/AR interaction standards lead to increased training efforts and make it difficult to change from one 3D or V/AR tool to another, regardless of whether this is within the framework of an existing work process or in the course of replacing the 3D or V/AR tool with another. The value added that can be achieved with V/AR is reduced, a lock-in effect can set in. Policymakers should make it their task to promote innovation measures and reduce barriers to competition. The barriers to the introduction, use and change of V/AR mentioned here must be counteracted. The most important thematic fields in V/AR standardization and -standardization were elaborated in the previous two chapters. Policy could promote the dissemination and intensity of use of V/AR technologies and methods through the following accompanying measures in the context of V/AR standardization:

1. <u>Promoting standardization activities in the V/AR context</u>
   As shown in the previous two chapters, there are still massive needs and activities in various standardization fields. Here, policy could support the ongoing activities organizationally, financially and in terms of marketing.
2. <u>Prefer standardized and normed solutions</u>
   Standardized, open V/AR solutions should be favored and demanded in public tenders and awards.
3. <u>Identify and occupy important, unoccupied standardization fields</u>
   The standardization of V/AR content and 3D data processing chains is an important lever for the economic success of V/AR use in companies. In the area of specific applications (vehicle development, machine development, Industry 4.0, medical technology, optics, active ingredient development, ...), there may well be thematically unoccupied areas in which standardization initiatives from Germany or Europe would have a realistic chance of gaining international acceptance. Here, politics could work together with science and associations to identify topics of great importance for Germany as a business location.
4. <u>Promote the involvement of domestic industry in V/AR standardization activities.</u>
   Domestic industry and science should be encouraged and, where appropriate, promoted by policy-makers to contribute domestic concerns to ongoing and future V/AR standardization activities/committees.
5. <u>Provision of information on V/AR standards and norms</u>
   The success of domestic V/AR solutions can be promoted by supporting existing and upcoming norms and standards. This requires transparent information on relevant existing

norms, standards and guidelines as well as on ongoing and planned initiatives. This position paper is already a first contribution to this.

B. Expectations to Industry

The economy, especially V/AR end users, can benefit massively from standardization. This can be a very direct, concrete benefit, such as keeping the option of switching open. But there are also indirect, more medium- to long-term benefits, such as contributions to a standard to which one's own solutions are already aligned. Industry could support V/AR standardization through the following flanking measures:

1. <u>Prefer standardized and normed solutions</u>
   In its own interest, the economy should also favor and demand standardized, open V/AR solutions in tenders.
2. <u>Identify and occupy important, unoccupied standardization fields</u>
   Companies are encouraged to look for "white spots" on the standardization map in their respective V/AR application areas or V/AR technology fields. If such a topic is found, it could be developed into a standardization initiative with partners - such as DIN or with competitors within the framework of a pre-competitive cooperation (coopetition).
3. <u>Engagement in V/AR standardization and standardization activities</u>
   Domestic industry should articulate its specific concerns in ongoing and future V/AR standardization activities/committees.
4. <u>Use existing V/AR standards and norms.</u>
   Companies should develop their V/AR solutions based on existing norms, standards and guidelines.

References

Boyce, Jill & Dore, Renaud & Dziembowski, Adrian & Fleureau, Julien & Jung, Joel & Kroon, Bart & Salahieh, Basel & Malamal Vadakital, Vinod & Yu, Lu. 2021. MPEG Immersive Video Coding Standard. Proceedings of the IEEE. PP. 1-16. https:/doi.org/10.1109/JPROC.2021.3062590

Boyce, Jill: MPEG-I Future Directions, SMPTE 2018, 2018, pp. 1-12, https:/doi.org/10.5594/M001830. Online: https://www.youtube.com/watch?v=DSAEpabeX78, retrieved 25.5.2021

Brennesholtz, M. 2018. 3 ‐ 1: Invited Paper: VR Standards and Guidelines. SID Symposium Digest of Technical Papers. 49. 1-4. https:/doi.org/10.1002/sdtp.12476.

Fraunhofer IPT: Standards für Augmented-Reality-Anwendungen in der Produktion. In: Virtual Reality Magazin; online unter. 2020, Aug. Vaterstetten, https://www.virtual-reality-magazin.de/standards-fuer-augmented-reality-anwendungen-in-der-produktion/ retrieved 25.5.2021,:

Green, L. 2017, May-June. IEEE Standards Association: Standardization Needs and Efforts in Virtual Reality/Augmented Reality; presentation at AWE USA, Santa Clara, California

McGarr, T., Boreham, M. 2018. The Requirement for Standards in the VR and AR Sectors. British Standards Institution & Futuresource Consulting Ltd, London, UK:

Mirt, J. 2016. Standardization Could Be A Major Problem For Virtual Reality; online: https://medium.com/@viarbox/standardization-could-be-a-major-problem-for-virtual-reality-9250efae1848, retrieved 25.5.2021

Open Geospatial Consortium: OGC Standards Roadmap, online: https://www.ogc.org/ roadmap, retrieved 25.5.2021

Perrey, C. 2019, Oct. Enterprise AR Interoperability and Standards Webinar. Augmented Reality Entreprise Alliance AREA. https://www.youtube.com/ watch?v=bj1yXUjfBOg, retrieved 25.5.2021

Preda, M., et al. 2019. ETSI GR ARF 001 V1.1.1. Group Report Augmented Reality Framework (ARF); AR standards landscape. ETSI Industry Specification Group (ISG), Sophia Antipolis, France.

Price, A. 2020. The role of international standards in virtual education and training systems. In: e-tech News & views from the IEC. International Electrotechnical Commission, Geneva/Switzerland.

Schreer, O. et al., 2020. XR4All Revised Landscape Report 2020. XR4All Consortium, Brussels, Belgium.

Shim, K.-W. 2019. Medical 3D Printing Scanning and Standards Requirements. In: ISO/IEC JTC 1/SC24 WG9 & Web3D Meetings, Seoul, Kore. https://www.web3d.org/event/web3d-korea-chapter-and-isoiec-jtc-1sc24-wg9, retrieved 25.5.2021

Stockhammer T. 2017. Qualcomm: MPEG Immersive Media. MP20 Standardization Roadmap.

Stockhammer, T. 2017. MPEG Immersive Media. MP20 Standardization Roadmap. In: International Telecommunication Union, ITU-T Study Group 16 Meeting, Geneva, Switzerland.

Timmerer, C. 2017, Dec. Immersive Media Delivery: Overview of Ongoing Standardization Activities. IEEE Communications Standards Magazine, New York, USA

Timmerer, C. 2018. MPEG Standardization Roadmap, ISO/IEC JTC1 SC29/WG11 MPEG, 122nd meeting, San Diego, CA, USA, Doc. N17506.

Verein Deutscher Ingenieure e.V.: VDI-Richtlinienarbeit. 2017. Grundsätze und Anleitungen. VDI 1000. VDI e.V., Düsseldorf.

Wajahat, A. K., Lee, S. R., & Kim, G. 2019. Standards for usability/UX evaluation process of VR/MAR contents. In: ISO/IEC JTC 1/SC24 WG9 & Web3D Meetings. https://www.web3d.org/event/web3d-korea-chapter-and-isoiec-jtc-1sc24-wg9, retrieved 25.5.2021, Seoul, Korea.

Wien, M., & Boyce, J. M. 2019. Stockhammer, Thomas; Peng, Wen-Hsiao: Standardization Status of Immersive Video Coding. In: IEEE Journal on Emerging and Selected Topics in Circuits and Systems 9 (1), 5 – 17

Won Lee, M., Suzuki, T., Overby, E., Whitlock, C., Minoo, K., … Cong, W. 2019. White Paper. Guidelines for Developing VR and AR Based Education and Training Systems. ISO/IEC JTC 1/VR AR for Education Ad Hoc Group; ISO/IEC.

Acknowledgements

The results shown have been achieved within the project "*Living Lab XR-Interact*", running from 2022 – 2024. This project is financed by the German Federal Ministry of Education and Research under the funding code 16SV8827.

Posters

Aggressive saliency-aware point cloud compression

Eleftheria Psatha, Dimitrios Laskos, Gerasimos Arvanitis,
and Konstantinos Moustakas

Department of Electrical and Computer Engineering, University of Patras, Greece

Corresponding authors: elefpsatha@gmail.com, dim-laskos@hotmail.com, arvanitis@ece.upatras.gr, moustakas@upatras.gr

Keywords: Point clouds compression, Saliency-aware quantization, Multi-saliency mapping

Introduction

Due to the increasing demand for accurate 3D scene representations in XR and immersive technologies, point clouds have become popular. However, quality point clouds require a lot of data, so compression methods are crucial. This paper presents a novel end-to-end compression scheme that utilizes geometry-based information and the user's position to achieve remarkable results for aggressive compression schemes with small bit rates. The main idea is to compress a simplified version of the original point cloud, that consists of different levels of resolution, highlighting the most visually significant parts of the point cloud. The compression scheme calculates four saliency maps based on the point cloud's geometry, distance from the user, visibility, and focus point. These maps are combined into a final saliency map "extended saliency" that quantizes different regions with a varying number of bits during the encoding process. The decoder reconstructs the point cloud using delta coordinates and solving a sparse linear system.

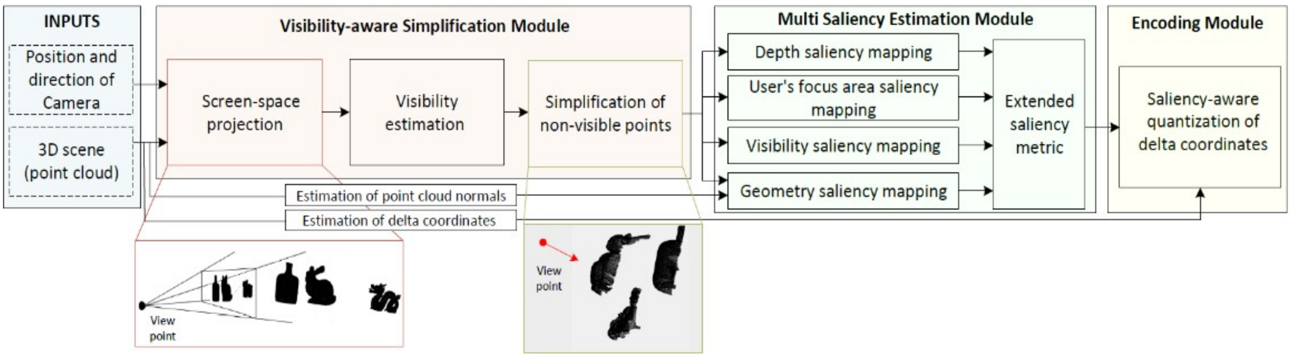


Figure 54. Pipeline of the proposed methodology. The Multi-saliency estimation module assigns a value to each visible point of the 3D scene indicating its perceptual significance. The encoding module compresses each point with a bit rate proportional to these values.

Saliency-aware Compression

A brief representation of the pipeline is illustrated in Figure 54. In a nutshell, the process starts by estimating the non-visible vertices of the point cloud based on the user's position and viewpoint. Then, the hidden vertices are simplified and four separate saliency maps are estimated, depending on: 1) Geometrical features, 2) Visibility, 3) Proximity between the user and the PC, and 4) User's focus point. Additionally, offline processes run only once for each point cloud, since they are independent of the location of the users and the direction of view. This includes i) the estimation of delta coordinates, based on the connectivity between the neighboring points, and ii) the estimation of point cloud normals.

Geometric Saliency. Motivated by the fact that geometric features, like high curvature regions, corners, and edges, usually convey important visual information, we estimate a saliency map based on the geometric importance of each point (Arvanitis et al., 2020). More specifically, we assume that areas with high-frequency spatial information are more perceptually significant and they must be preserved in contrast to flat areas. In our proposed geometric saliency scheme, we combine an eigenvalue-based step, which extracts saliency features by decomposing local covariance matrices defined in small regions around each point (Figure 55 (a)).

Visibility Saliency. The used visibility operator is a metric defining the confidence that a point is not occluded. More specifically, the closer to 1 is, the more certain we are that the point is visible. We create a saliency map based on this operator, assuming that points with high values are visually more salient (Figure 55 (b)).

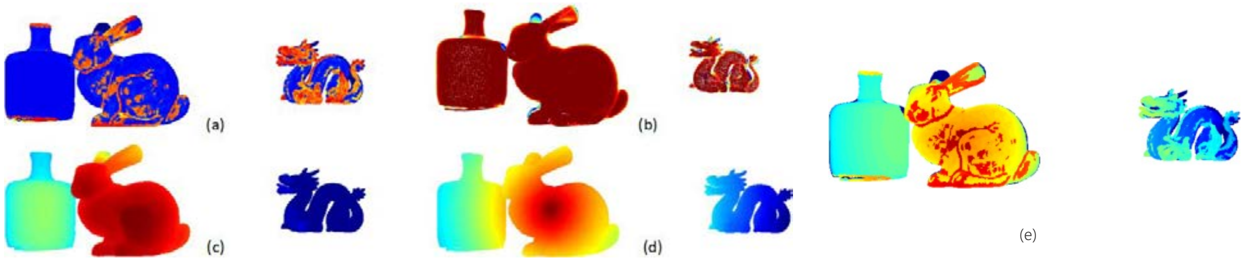


Figure 55. Heatmap visualization of the saliency map based on (a) vertices geometry, (b) vertices visibility, (c) depth, (d) the users' focus, (e) Heatmap visualization of the final saliency map.

Depth from Viewpoint Saliency. Due to the restricted depth of field of the human eye, objects that are far from the viewpoint tend to appear blurred. In detail, visual acuity is reduced at the parts of the field of view with higher depth values. Motivated by this fact, we calculate a saliency map based on depth difference from the point of view (Lee et al., 2008). The user will not notice the visual difference in quality reduction (Figure 55 (c)).

Focus Point Saliency. The point that the users' eyes are looking at is called as "focus" point and can be derived from their position and viewing direction. It is known that humans perceive more details close to the focus point than at its periphery. We estimate a saliency map by incorporating a peripheral blur effect that progressively reduces the quality of the scene at points located at a certain distance focus point. The quality reduction in the periphery adds a realistic sensation and is not noticeable by the user (Figure 55 (d)).

Based on a combination of these maps, each point of the simplified point cloud is associated with a value that indicates its significance (Figure 55 (e)). Afterwards, the point-cloud delta coordinates, which are calculated using an approximate connectivity graph, are scaled, quantized and eventually entropy coded. Each point is encoded with a different number of bits depending on its extended saliency. Finally, on the decoder side, we reconstruct the point cloud using the decoded delta coordinates.

Experimental Results and Analysis - Qualitative Evaluation

In Figure 58, we present the reconstructed results of our approach in comparison with the results of the other methods. Since our proposed scheme compresses only geometry data, we have given the original colors to the reconstructed models for visualization purposes. We also provide enlarged details of each one of the reconstructed models, for easier comparison among the methods and the corresponding bpp ratio. Additionally, we illustrate the heatmap visualization of each reconstructed model, highlighting in different colors (i.e., red color represents small differences while blue represents big differences) the Euclidean distance of each vertex with the corresponding vertex of the original model, and the mean Euclidean distance. It is evident that our method preserves details in geometry for very small bit rates and provides good visual accuracy compared to G-PCC's octree and trisoup[11] where the quality of reconstructed models is relatively poor, leading to high quantization errors. We further extend our studies by examining the performance of our proposed compression scheme in areas of distinct perceptual significance. To that end, we separated the input scene into 4 layers based on the extended saliency metric (Figure 57). Therefore, each layer is encoded with a different number of bits. In Figure 56 the rate distortion curves for each layer is presented. As we can observe, layer 4 is almost unaffected by decreasing the bit rate, since it consists of points encoded with zero bits [3]. On the other hand, layer 1 exhibits higher levels of distortion. This occurs because the areas with high geometric saliency represent high frequency information. In Table 5 we compare our saliency-aware encoding method with a scheme that assigns the same number of bits to each layer. By assigning more bits to layer 1, our method exploits the aforementioned characteristic and achieves better reconstruction results.

Conclusion

In this work, we presented a geometry-based, end-to-end compression scheme for large point clouds scenes. Our proposed method highlights the most visually significant parts of the point cloud and compresses the position of each point based on an extended saliency metric that combines viewer's relative position and geometric saliency. The quality reduction in perceptually insignificant parts of the scene adds a realistic sensation and is not noticeable by the user, even for aggressive compression rates. Qualitative tests have shown that the quality of the reconstructed point clouds remains almost unaffected even for very small bit rates. We could extend our studies by integrating the saliency-aware encoding scheme to user- interactive XR rendering applications in order to increase compression efficiency in such scenes.

Table 5. *D1 comparisons at 1.2 bpp.*

Quantization	L1	L2	L3	L4
Saliency	49.18	55.75	54.51	53.14
Uniform	46.10	53.26	53.23	53.10

[11] https://mpeg-pcc.org/

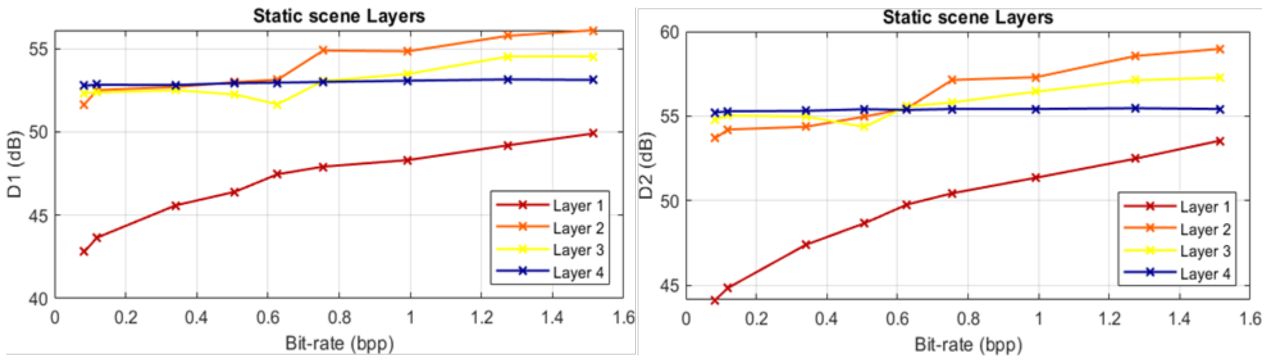


Figure 56. PSNR comparisons for different layers.

Figure 57.Saliency-based layer separation.

Figure 58. Indicative illustrations of (a) original models, and reconstructed results, heatmap visualizations of the Euclidean distance between the original and reconstructed models along with the mean Euclidean distance, using: (b) the G-PCC octree approach, (c) the G-PCC trisoup approach, (d) our approach.

References

3DG Group. 2019. Common test conditions for point cloud compression. ISO/IEC JTC1/SC29/WG11 Doc. N18474.

Arvanitis, G., Lalos, A. S., & Moustakas, K. 2020. Robust and fast 3-D saliency mapping for industrial modeling applications. IEEE Transactions on Industrial Informatics, 17(2), 1307-1317. https://doi.org/10.1109/TII.2020.3003455

Lee, S., Kim, G. J., & Choi, S. 2008. Real-time tracking of visually attended objects in virtual environments and its application to LOD. IEEE Transactions on Visualization and Computer Graphics, 15(1), 6-19. https://doi.org/10.1109/TVCG.2008.82

Enhancing health and safety monitoring of construction sites with Mixed Reality

Thomas Papaioannou, Tina Katika, Fotios K. Konstantinidis, Konstantinos Routsis, Spyridon Nektarios Bolierakis, Georgios Tsimiklis, Giannis Karaseitanidis, and Angelos Amditis

Institute of Communication and Computer Systems (ICCS), Athens, Greece
Corresponding author: anastasovitis@iti.gr

Keywords: Virtual museum, serious game, storyliving, cultural heritage, virtual reality.

Introduction

Mixed Reality (MR) is an innovative technological paradigm that exhibits considerable potential in revolutionizing health and safety monitoring practices and emerges as a pivotal tool for promoting safety inside work environments. Within industrial, construction, and manufacturing settings, MR effectively facilitates the visualization and assessment of potential hazards in authentic environments, thereby empowering workers to efficiently discern and mitigate risks. Notably, workers equipped with MR headsets can readily receive visual cues when approaching hazardous areas or encountering unsafe conditions (Katika et al., 2022). This exceptional capability substantially enhances safety measures, empowering remote experts to identify latent hazards, provide immediate guidance and support to on-site workers through intricate procedures, and ensure compliance with safety protocols (Moore & Gheisari, 2019). Researchers have recognized the significant potential of MR in elevating diverse facets of construction processes within construction sites. For safety training and visualizing safety information, MR may improve safety awareness and hazard identification on construction sites (Wolf et al., 2022). At the same time, integration of indoor positioning systems (IPS) and building information modeling (BIM) enhances situational awareness in construction and such visualizations in MR enable efficient visual management, improve planning and productivity (Reinbold et al., 2019).

Problem statement

MR enhances worker safety training by creating immersive environments, visualizing real-time safety information, and connecting to IoT sensors and wearables. Integrating MR applications can result in a safer working environment, increased worker awareness, improved training efficiency and better safety standards compliance. The current study prioritized the design and development of an MR application with specific functionalities: 1) enhanced visualizations for better understanding of potential dangers and safety procedures, leading to improved situational awareness; 2) real-time data monitoring for proactive monitoring, early hazard detection, and timely interventions; 3) documentation and reporting for simplifying incident reporting and safety inspections; and iv) collection of safety performance analytics for gaining insights on risk factors and improving overall safety.

IoT ecosystem

The MR application was developed to target HoloLens 2 devices and was used by the construction site's health and safety officer (HSO). To provide a clearer understanding of the MR application, it is essential to contextualize its role within the construction site. An IoT ecosystem, consisting of various components, called enablers, was constructed inside the construction site, ensuring secure collection and visualization of all the generated flowing data. The key to effectively integrating the MR component lay in harnessing the existing IoT devices of the ecosystem and their benefits. This included a comprehensive process of identifying the IoT devices that were relevant to the MR enabler, encompassing an array of sensors, actuators, smart objects, and wearables, all actively gathering or generating data (See below Figure 3b). The equipment was designed to be both light and comfortable, ensuring minimal disruption to the workers' tasks. The data was transmitted to relevant enablers via an Edge Data Broker (EDB), an enabler responsible for the communication between the enablers, using MQTT protocols. After processing, the enablers promptly activated alerts in the event of emergencies, disseminating only essential information to relevant enablers, with utmost regard for the workers' privacy. Those IoT devices are parts of other enablers, that coexist with the MR inside the ecosystem. Last but not least, a configuration layer was crucial to be designed, in order to ensure that the MR enabler could stay up-to-date and receive data from the rest of the enablers at all times.

MR enabler: Building Information Modeling

The MR enabler significantly transforms construction site management by offering a powerful platform to observe the digital twin of the site, effectively identify its dangerous zones (Figure 59), and analyze efficient evacuation routes (Figure 60). Moreover, it provides a comprehensive graphical representation for individual workers within the construction site. The BIM model of the construction site is designed to accommodate updates and can be uploaded to the corresponding enabler. During runtime, the MR enabler receives the updated BIM model through REST protocols, ensuring real-time access to the latest project information. The live location of the workers on-site is being sent by the EDB.

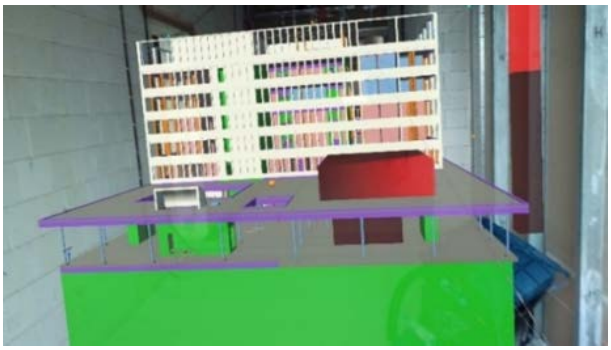


Figure 59. The dangerous zones of the site

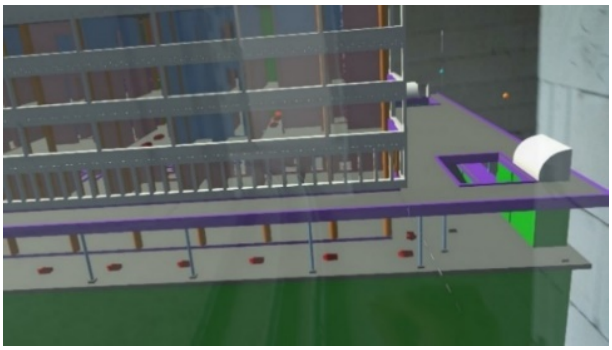


Figure 60. The evacuation routes of the site

MR enabler: Alert system

The application excels at promptly notifying the HSO whenever a worker requires assistance, providing comprehensive and essential information. This data encompasses the worker's precise location and any specific health conditions recorded by the corresponding enablers and sent to the MR through the EDB (Figure 61 & Figure 62). To provide the workers' location points and facilitate seamless navigation for the HSO within the site (Figure 63 & Figure 64), the application leverages an indoor localization system based on Ultra-Wideband technology (UWB).

Figure 61. A new alert was generated for a construction worker.

Figure 62. Worker's heartbeat, tracked by their wearable

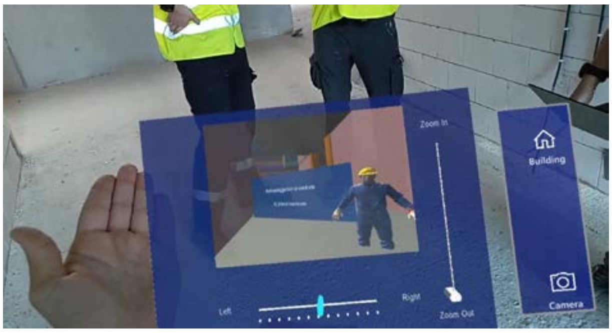


Figure 63. Worker's precise location inside the building model

Figure 64. A special wearable is worn by the site's workers that tracks their location and their heartbeat at all times

MR enabler: Documentation and reports

At any given moment, the HSO can readily access real-time information on all the workers that are currently inside the construction site, allowing for thorough examination of their data (Figure 65). Moreover, the application grants the capability to generate new reports in the event of non-compliance incidents (Figure 66). The report may include textual or multimedia information to provide a comprehensive summary of the incident. The MR application establishes communication with the relevant enabler, facilitating the exchange of worker's documents and reports, all accomplished through the utilization of REST protocols.

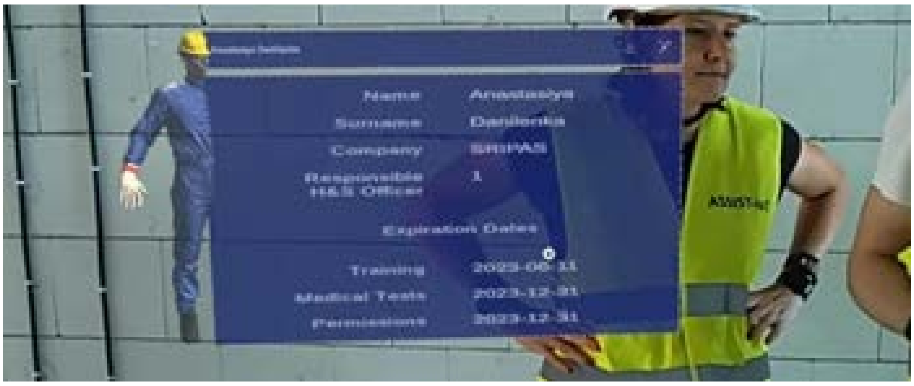


Figure 65. *Inspection of the construction site worker's information*

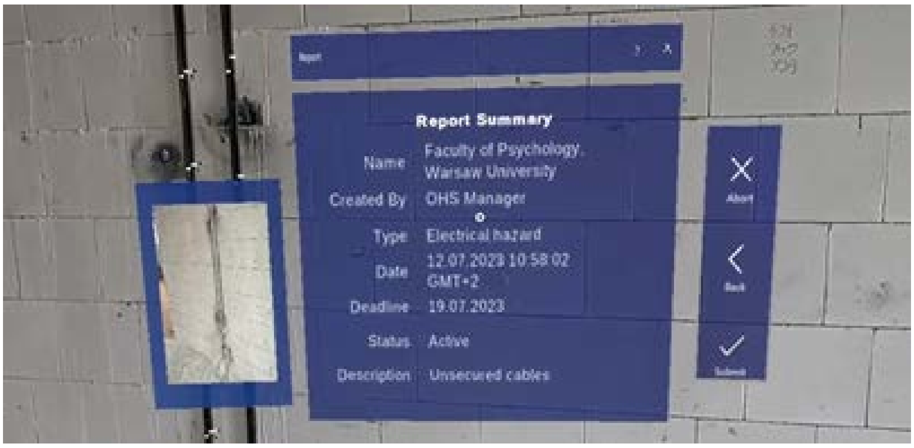


Figure 66. *Generate and store a new report through the MR enabler*

Acknowledgement

The research of the MR enabler was part of the Assist-IoT project, which has received funding from the European Union's Horizon 2020 Research and Innovation program under the Grant Agreement No. 957258. The opinions expressed herein are those of the authors and the European Commission is not responsible for any use that may be made of the information it contains.

References

Katika, T., Konstantinidis, F.K., Papaioannou, T., Dadoukis, A., Bolierakis, S.N., Tsimiklis, G., & Amditis, A. 2022. Exploiting Mixed Reality in a Next-Generation IoT ecosystem of a construction site. IEEE Int. Conf. Imaging Systems and Techniques (IST), pp. 1-6. https://doi.org/0.1109/IST55454.2022.9827726

Moore, H.F. and Gheisari, M. 2019. A review of virtual and mixed reality applications in construction safety literature. Safety, 5(3), 51. https://doi.org/10.3390/safety5030051

Reinbold, A., Seppänen, O., Peltokorpi, A., Singh, V., Dror, E. 2019. Integrating Indoor Positioning Systems and Bim To Improve Situational Awareness. Proc. 27th Annual Conference of the International Group for Lean Construction (IGLC). https://doi.org/10.24928/2019/0153 .

Wolf, M., Teizer, J., Wolf, B., Bükrü, S. and Solberg, A. 2022. Investigating hazard recognition in augmented virtuality for personalized feedback in construction safety education and training. Advanced Engineering Informatics, 51, p.101469. https://doi.org/10.1016/j.aei.2021.101469

AEOLIAN AR mobile application: Disaster tales to enhance preparedness of citizens and bridge the Risk Perception Action Gap

Orestis Sampson, Panagiotis Michalis, Chrysoula Papathanasiou, Giannis Karaseitanidis, and Angelos Amditis

Institute of Communication and Computer Systems, Athens, Greece

Corresponding author: orestis.sampson@iccs.gr

Keywords: Crowdsourcing, AR Campaigns, Risk Information, Disaster Tales, Risk Awareness, Climatic Hazards, Critical Key Locations, AR Virtual Learning, Immersion, User Engagement

Introduction

The impact of climate change has become a significant global challenge, leading to an increase in climatic risks. As climate change progresses, weather events are expected to become more severe (Michalis et al, 2022a) posing a considerable threat to critical assets located near watercourses and vulnerable societal functions susceptible to extreme droughts and floods. Traditionally, resilience-building approaches have focused on managerial and technical crisis responses, while citizen engagement regarding risks and preparedness measures has been predominantly one-way and top-down. However, research has shown that this top-down approach is insufficient (Michalis et al, 2022b), prompting a need for more active citizen involvement in disaster management through a citizen-oriented, bottom-up approach.

Related work and problem statement

To effectively adapt to current and future hazardous events, leveraging emerging technologies to support disaster preparedness and response is essential (Batzos et al, 2023). Despite the potential of new technologies and other efforts in response training, their effective implementation in climatic disaster risk management phases has not been fully explored. Other efforts have been made (be-alert, Padovapartecipa, protezione-civile (Regione Veneto), GI-Polis) to empower user response and participation for disaster resiliency for citizens but have either been limited to be operational in a defined geographical area or are not designed to support bi-directional communication between citizens and authorities or both.

To address the lack of citizen engagement during disaster preparedness and response phases of evolving climatic risk situations, this study introduces the AEOLIAN system. The proposed system employs augmented reality (AR) technology within a mobile application to merge real environments with virtual objects, offering disaster tale narratives in a user-friendly and easily digestible format. At the same time the AEOLIAN system provides a user-friendly interface to the authorities to effectively organize training/educational activities and communicate with the citizens by disseminating useful guidelines.

The AEOLIAN system

The AEOLIAN system modules contain a comprehensive set of tools for the prompt and efficient reporting of incidents and observations. The modules also allow for effective communication with citizens, ensuring that all relevant information is gathered and reported. Finally, the data that is collected can be visualized on a map layer, providing a clear and concise overview of the situation at hand. The logical architecture consists of the modules presented in Figure 67.

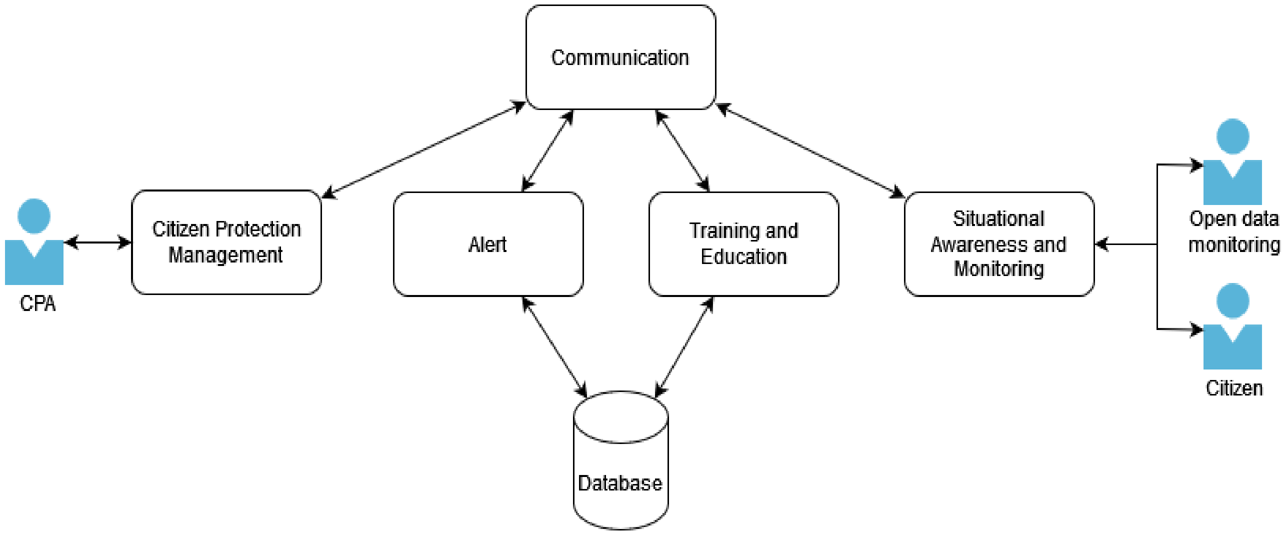


Figure 67. Logical Architecture of the AEOLIAN system

The system comprises two primary components catering to two main user groups. The first component is the Content Management System (CMS), a web-based application that allows Citizen Protection Authorities (CPA) to manage and oversee the content uploaded by mobile users. The CMS is comprised of two distinct parts: the back-end, which stores resources relevant to campaigns, such as files, metadata, and locations, in a database and provides APIs; and the front-end, which retrieves these resources through APIs to display the CMS interface in the browser. Additionally, CPAs can create engaging AR campaigns accessible to end-users through the mobile application (Figure 68).

The second component is the AEOLIAN AR mobile application, primarily serving as a platform for end-users to access AR campaigns while also offering other functionalities. It offers various features to promote engagement and encourage citizens to familiarize themselves with safety practices and historical incidents through multimedia and text guidelines. These functionalities include AR visualization of CPA-provided material, gamification features for enhanced user engagement, and the creation of hazard campaigns by CPAs focused on specific risks and natural disasters. By utilizing these features, citizens can better understand potential risks in various scenarios and learn how to respond effectively.

Each AR campaign uploaded to the CMS contains all necessary information to create an engaging AR experience. Campaigns have specific goals and objectives and include a Point of Interest (PoI), which can be a location or area on a map containing media and educational content for mobile users. The CMS enables administrators to perform various actions for each AR campaign, such as adding the PoI's location, providing title and description in text format, and attaching media content like images and videos.

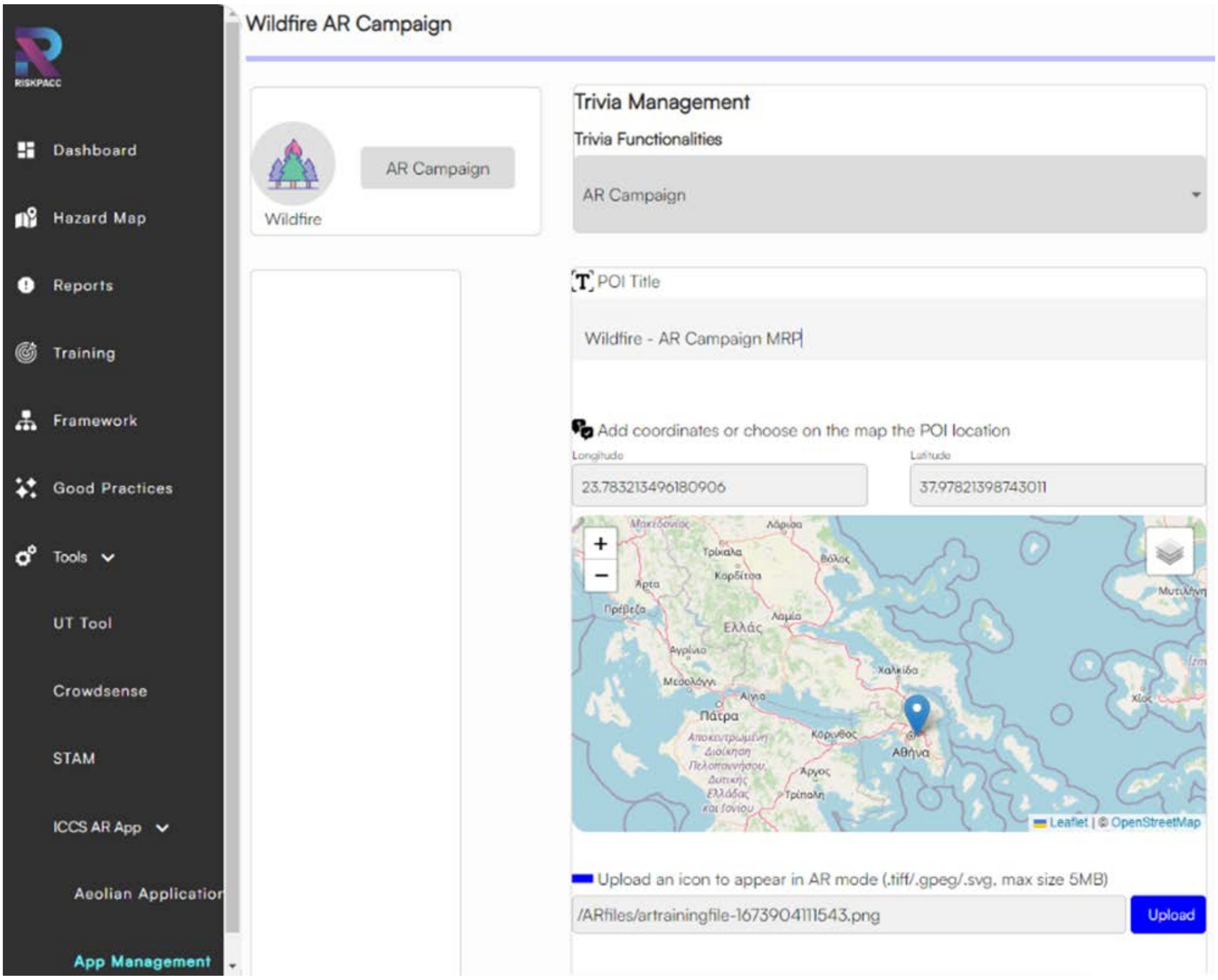


Figure 68. *AR campaign creation in the AEOLIAN CMS*

AEOLIAN mobile application core components

The AEOLIAN application was developed through iterative co-creation, incorporating feedback from CPAs, citizens, and volunteers. During the initial design phase, the tool aimed to create a solution supporting various phases of the disaster management cycle, with a primary focus on prevention, preparedness, and response to natural hazards. The core functionalities, including AR, were collaboratively discussed with CPAs from different case studies. The features were first determined and later refined based on targeted interactions with end-users, considering their needs and requirements. This approach ensured the application's relevance and effectiveness for those who will ultimately use it.

The AEOLIAN AR mobile application was developed in the Unity game engine[12] using AR Foundation and Mapbox. The app is compatible with Android smartphones and requires Global Positioning System (GPS) to provide localization information. To activate the AR experiences, the application fetches data generated by the CMS during the AR campaign creation process through a

[12] Unity Game Engine: https://unity.com/

REST API. For users to properly utilize the app, they must allow the use of the camera and location settings, as well as have an internet connection.

Currently the Aeolian AR mobile application has reached its almost-final stage. The application has five core functionalities: 1) Hazard Map, that presents an interactive map with pinpoints that correspond to natural or man-made hazard events that have occurred in an area and have been posted by local CPAs for awareness raising and information sharing purposes, 2) Reporting, that allows users to create a report during an event and communicate it to the CPAs and optionally to directly interact with them via a chatting functionality, 3) Training that is classified into Disaster Training (which includes quiz questions and tasks), AR campaigns (which enables training of users through an integrated AR functionality) and Good to know (which is a repository with documentation on disaster risk management), iv) Notifications that are updated when a hazard event is updated or added, a training session is added or the status of a report changes and v) Emergency call that enables the user to directly dial the European Emergency Number 112 (Michalis et al, 2023). The mobile application allows for the visualization of virtual content at specific locations after creating AR campaigns through the CMS administrator and adding the corresponding content. By leveraging the familiarity of citizens with certain places, the awareness of these key locations can be increased, which can prove crucial during emergencies. The use of AR campaigns can also serve as a tool for presenting disaster tales through a narrative approach that is more accessible and attractive to the public (Papathanasiou et al, 2023). Rather than using "static" risk mapping, a story-line approach can better represent the sequence of events during a disaster and recovery phase.

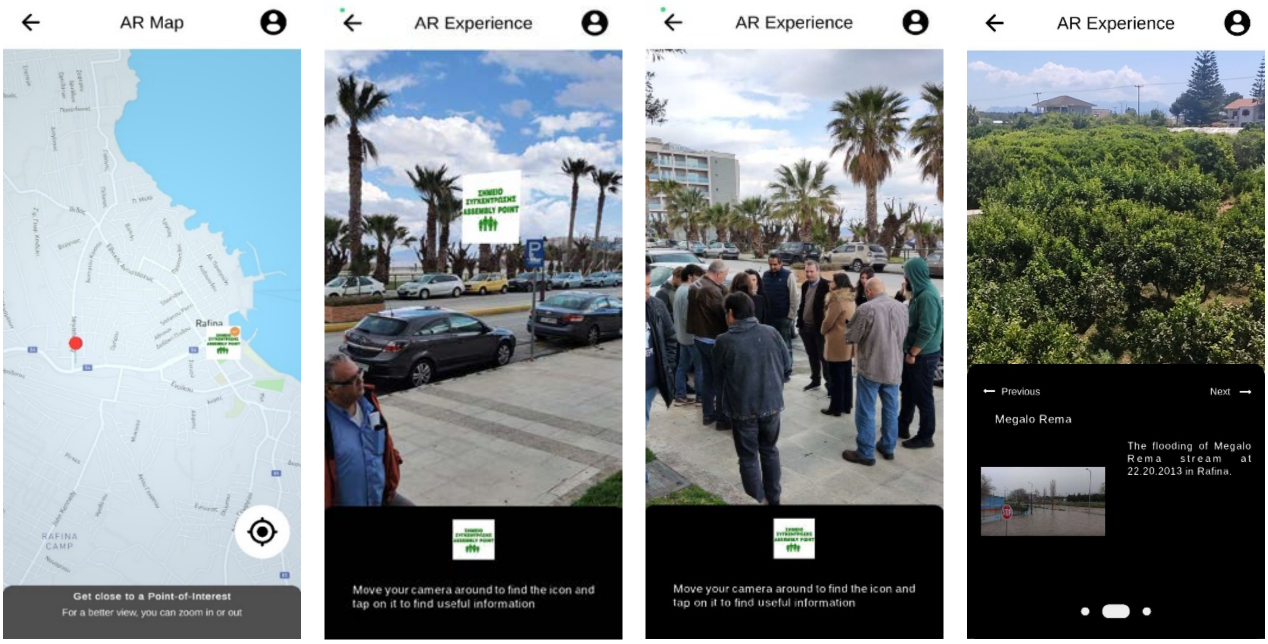


Figure 69. Four UIs of the AEOLIAN app: a) AR map to help the user get to the correct location, b) AR icon related to the Assembly Point, c) Crowd gathers around the Assembly Point during demonstration, d) Description and content for the disaster tale in the area.

Aiming to cover the need for enhanced awareness on different hazards, the AR functionality of the AEOLIAN application includes disaster tales that can be created for any type of hazard, natural or man-made that is of interest to the citizens in that area, either because of its frequency of occurrence or because of its impact. Targeted information for specific events can be incorporated in the disaster tales in an area that is particularly prone to these specific types of disaster, illustrated with relevant pictures and videos, presented in Figure 69. This information may include impacts of historic events in the area, statistics, lessons learnt (thus expanding the applicability of the app to the recovery phase of disaster management). By following the campaigns with the disaster tales, users become aware of the severity and the impacts of these hazards in their area, and such awareness raising fosters in its turn the bridging of the Risk Perception Action Gap (RPAG).

Further to that, the AEOLIAN application provides the option to keep the citizens informed about critical key locations in their area, through appropriately designed AR Campaigns. More specifically, CPAs can design an AR Campaign informing the users about safe public spaces, away from dangerous zones, where citizens can find assistance (e.g., water, food, first aid, heated rooms) in case of emergency. An AR Campaign to provide response information was designed to provide useful information about an open space assembly point where citizens can be gathered in case of a wildfire event. Once users are on-site at the defined locations for each campaign, they can be navigated to the campaign and retrieve additional information either for the relevant hazard or for the site per se (see Figure 69).

Conclusions

The AEOLIAN system offers potential benefits that go beyond its immediate application. By improving understanding and communication of climatic risks, more people can become aware of the dangers posed by natural hazards, prompting them to take anticipatory measures and enhance preparedness. Future work will investigate the integration of chat bots and smart agents to enhance the AR experience and handle some parts of the communication pipeline.

Acknowledgement

This research has been financed by the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 101019707, project RiskPACC (Integrating Risk Perception and Action to enhance Civil protection-Citizen interaction).

References

Batzos, Z., Saoulidis, T., Margounakis, D., Fountoukidis, E., Grigoriou, E., Moukoulis, A., ... & Filippidis, A 2023. Gamification and Serious Games for Cybersecurity Awareness and First Responders Training: An overview. https://doi.org/10.36227/techrxiv.22650952.v1

BE-Alert, Belgium's National Crisis Center pilot project for emergency notifications: https://www.be-alert.be/en

GI-Polis, Android mobile application for the Local Authorities to send Notifications to the Citizen devices, based on their location: https://play.google.com/store/apps/details?id=com.geoinformaticspcc.gipolis.android&hl=el&pli=1

Michalis, P. & Vintzileou, E. 2022. The Growing Infrastructure Crisis: The Challenge of Scour Risk Assessment and the Development of a New Sensing System Infrastructures, 7, 68. https://doi.org/10.3390/infrastructures7050068

Michalis, P., Anniés, J., Papathanasiou, C., Ouzounoglou, E. & Admitis, A. 2022. Co-creation as an approach to bridge the Risk Perception Action Gap and customise crowdsourcing tools to respond effectively to climatic risks. 3rd Int. Conf. Natural Hazards & Infrastructure, ICONHIC 2022, 5-7 July 2022, Athens, Greece

Michalis, P., Georgiou, K., Sampson, O., Nousis, V., Palazis, M., ... & Amditis, A. 2023. AEOLIAN: A new Augmented Reality mobile application to enhance disaster training and response to evolving hazards. EGU General Assembly 2023, Vienna, Austria, 24–28.4. EGU23-1073. https://doi.org/10.5194/egusphere-egu23-1073.

Padovapartecipa, Web application for reporting in the Municipality of Padova: https://padovapartecipa.it/

Papathanasiou, C., Michalis, P., Stavrou, K., Tsougiannis, E., Anniés, J., ... & Amditis, A. 2023 Enhancement of local community resilience to natural and man-made disasters through the application of co-created novel technological tools. In: EGU General Assembly 2023, Vienna, Austria, 24–28 Apr 2023, EGU23-1381. https://doi.org/10.5194/egusphere-egu23-1381

Regione Veneto, Web application for Civil Protection Alerts: https://www.regione.veneto.it/web/protezione-civile/cfd

Augmented Swiss Heritage: Visualizing cultural heritage with Augmented Reality

Nadine Ganz, Simonne Bosiers, and Onna Rageth

University of Applied Sciences of the Grisons, Chur, Switzerland

Corresponding author: *nadine.ganz@fhgr.ch*

Keywords: Augmented Reality, Cultural Heritage, Tourism, Immersion

Introduction

Swiss tourism still has great potential in terms of digitalization. The current tourism offer is often reduced to the beautiful landscape with its mountains and lakes. Apart from nature, Switzerland offers a lot of culture and history. To make the cultural heritage accessible for the tourists, we developed an Augmented Reality (AR) app for the destination Davos in Switzerland. The app offers four routes, which lead the visitors through Davos and its surroundings. The sights along the routes are enriched with AR content, like animation, videos, or interactive puzzles.

The goal of the project was to investigate how AR can be used to visualize the cultural heritage of a city in an immersive and participatory way. Our app should serve as a model that can be multiplied and used by other tourism organizations and destinations. Furthermore, we elaborated and defined the necessary steps of the production process to deliver a high-quality result in a cost-effective and timely manner.

Methodology

The production process of the app consisted of nine steps, as shown in Figure 70. At the beginning, the whole team went with the project partners to Davos and visited the sights and museums of the city. A ranking of the most important sights has been made. Afterwards, we developed different personas as well as user scenarios, and defined the locations for the AR content in Davos. In the next phases, the content creation and implementation of the app took place. A detailed description of that can be found in the next sections. The design and implementation of the app was an iterative process, consisting of prototyping, user testing, programming and UX designing.

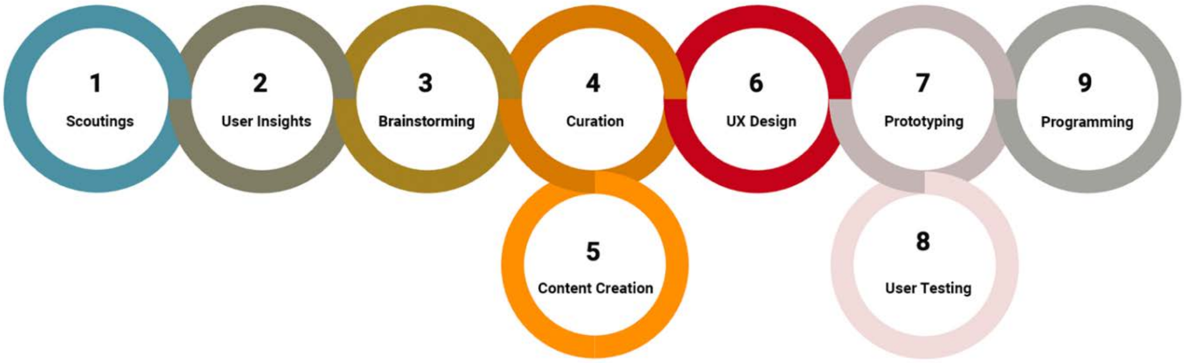


Figure 70. *Steps of production process*

Content Creation

The content creation is an important part of the production process because the content of an app is crucial for its usefulness and the satisfaction of the users. In advance of this phase the content needs to be prepared. The curation of historical artefacts involved difficulties, because only a small number of historical artefacts is digitalized. It was challenging to find originals of the artefacts, especially in high resolution form. Therefore, the AR content needed to be adapted to the available materials. The curation was an iterative process in collaboration with the museums and software developers, keeping the financial and technological limitations in mind.

AR is suited to make things visible that cannot be experienced in the real world, e.g., historical artefacts like famous artwork, which are not materially present at the place of origin. Historical events and figures can be brought to life through digitally supported storytelling. For our app we designed sixteen different AR pattern that can be replicated in form and interactivity with other content. They are equipped with different types of media and interactivity to offer users a wide range of experiences. These include interactive 2D paintings of the artist *E.L. Kirchner*, audios of historical figures giving a fictional monologue, and animations. In terms of content, the AR stations are based on the exhibition catalog (Hess, 2021) from the *Germanischen Nationalmuseum* in Germany.

Implementation

The app was implemented for mobile devices with the game engine *Unity* and the AR framework *AR Foundation*. The map of Davos was built with *Mapbox*, a maps SDK for *Unity*. For the User Interface (UI), we used the standard 2D *Unity* UI components. Our UI has a clear, concise structure and a barrier-free design for people with disabilities. To make the interaction with the AR content easier, we display hints like text instructions or a circle for placing objects.

The AR content resides on *discover.swiss*, an open backend service platform based on *Microsoft Azure Spatial Anchors* for locating the AR models and a cloud for retrieving them. Every AR content is linked to a specific location in Davos. When the user is on site at the AR station, the app recognizes the location and loads the AR content from the cloud. Beforehand, our team had to define all the positions of the AR content on site by setting so-called anchors with our self-developed anchor app. Thereby, the anchor app collects feature points from the environment and generates a 3D point cloud of the location. Additional information like GPS is saved, too. The anchor app publishes the anchor metadata (i.e., 3D point cloud) to our AR content service. The 3D models of the sights are organized with the Addressable Asset System from *Unity*. In our case, the assets are prefabs with all *Unity* components and scripts. Each sight is represented as one asset. They are bundled into one catalog, which is uploaded to our AR content service. When selecting the locations, certain aspects must be considered. The environment should be constant and not changing, hence the anchors should not be placed on moveable items. Permanent objects like walls or statues are well suited for anchors. Small changes in the environment can make the recognition of the location harder, e.g., the blinds at a house were closed while setting the anchor and open during recognition. Apart from technical requirements, the location should be at a reasonable distance from large crowds or gatherings of people and must be easily accessible under various weather conditions.

Results

We developed an AR app for the visitors of Davos to experience the cultural heritage of the city in an immersive and participatory way. On four geographically routes through Davos, the user can take a journey into the past.

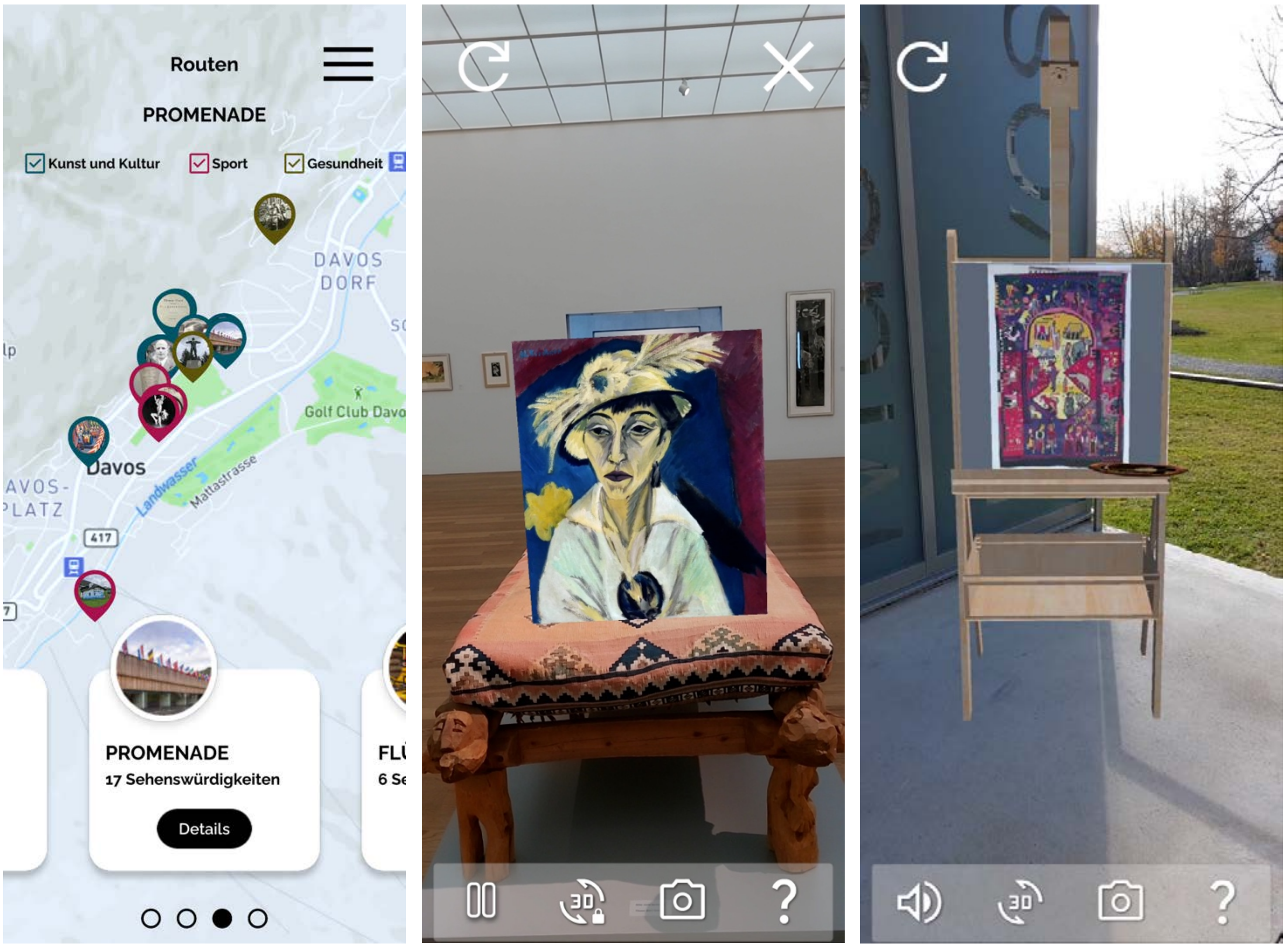


Figure 71. *Screenshots from App: Map of Davos with routes and sights (left), augmented historical artefact in the Kirchner Museum (middle), virtual 3D easel with famous painting of E. L. Kirchner (right).*

The app contains 38 sights and covers three thematic areas: art and culture, health, and sports. Each sight is accompanied by text and audio that provides the user with its history. The sights differ in their form and interactivity. Figure 71 shows two examples. In the middle the famous *Bed for Erna*, a historical artefact of the *Kirchner Museum*, can be seen. *Erna Schilling* was the partner of the artist *E.L. Kirchner*. To know the story of the bed, we enriched it with a portrait of *Erna* and brought her to life with an animation where she tells the audience the history of the bed. At the right in Figure 71 a virtual 3D easel with a historical tapestry from *E.L. Kirchner* is placed in the world. At the beginning, the tapestry is black and white. The task of the user is to weave the tapestry in color while he listens to the story of the artefact.

The best experience of the app is on site in Davos, following the four different routes using the map shown in Figure 55 on the left. However, there is also the possibility to view the AR content at home. In this case, the users have to place the content manually in their environment. The app can be downloaded via App Stores, also there is a cinematic summary available to get a first insight into the app (FHGR, 2022).

Conclusion

In our project we developed an AR app to visualize the culture and history of the city Davos. The app should serve as a showcase and is replicable to other tourist destinations in Switzerland. Our developed AR patterns show different forms of how historical sights can be augmented in an interactive and entertaining way. For the future, it is planned to create further AR experiences for other tourist destinations in Switzerland based on the presented app.

Acknowledgements

This work has been funded by the Innotour Program of the State Secretariat for Economy (SECO, Switzerland) and Kreativfonds Davos. Additional thanks to our project partners: Kirchner Museum Davos, Heimatmuseum Davos, Afca AG, discover.swiss and Destination Davos Klosters.

References

Fachhochschule Graubünden (FHGR). 2022. Augmented Swiss Heritage. https://www.youtube.com/watch?v=J3tr-jSbbTw

Hess, D. 2021. Europa auf Kur. Ernst Ludwig Kirchner, Thomas Mann und der Mythos Davos. Germanisches Nationalmuseum Abt. Verlag.

Towards an Immersive WebXR-based Solution for Smart Farming: Enhancing Transparency and Comprehensibility in Agricultural Management

Irene Bouzón, Ignacio Pedrosa, Moisés Muñiz, and Roberto Abad

Technological Center for Information and Communication (CTIC), Gijón, Spain

Corresponding author: irene.bouzon@fundacionctic.org, ignacio.pedrosa@fundacionctic.org, moises.muniz@fundacionctic.org, roberto.abad@fundacionctic.org

Keywords: WebXR, Virtual Assistant, Smart Farming, UX, Comprehensibility, Sustainability, AI

Introduction

Rapid growth of Artificial Intelligence (AI) in recent times has seen an increasing adoption of machine learning techniques in various research areas. As a result, AI technologies have found their way into everyday life, impacting various user segments positively. Despite these advancements, the complexity of underlying algorithms poses challenges for end-users, especially those with limited digital proficiency, such as farmers. Consequently, transparency and comprehensibility issues arise, impeding human-computer interaction for end-users. This situation calls for innovative solutions that cater to user needs and enhance their understanding of AI-driven systems.

VR and AR hardware have witnessed significant growth in the consumer market, providing immersive computing experiences with new opportunities and challenges. Ensuring the seamless integration of the web in this environment is crucial for a first-class user experience. The WebXR Device API offers essential interfaces to developers, enabling them to build compelling, comfortable, and secure immersive applications across diverse hardware devices. This standard facilitates rendering 3D scenes in AR or VR, encompassing various web-connected devices, while effectively handling distinct user input controls and output display capabilities.

Furthermore, with the global population expected to increase substantially by 2050, agricultural production faces a tremendous surge in demand. Developing cost-effective and sustainable agricultural solutions becomes imperative, requiring a 90% intensification in technological farming to meet market demands while adhering to environmental and *twin transition* objectives. To address these challenges, an AI-based approach is needed to create a smart farming environment that offers real-time monitoring and management of agricultural issues.

This paper proposes an integral solution based on WebXR standards that aims to provide an immersive and trustworthy environment for farmers. Guided by a virtual agent, this solution seeks to offer real-time understandable information to farmers, empowering them to automate processes and make data-driven decisions to optimize production while adhering to sustainability standards. By fostering transparency and user-friendliness, this approach aims to **bridge the gap between AI technologies and end-users**, particularly in the agricultural sector.

Methods

To achieve the defined objectives, a **digital twin of an agricultural scenario** has been built and deployed as an immersive platform, compliant with **WebXR** standards in order to provide accessibility and ease of use. This digital twin graphically represents the agrarian landscape as a three-dimensional environment based on real world geographic data derived from **satellite imagery**, ensuring a high level of accuracy and authenticity in the virtual representation. The premises depicted in this area include a distinctive infrastructure such as a climate simulator that houses multiple terraces containing a variety of crops (i.e. faba beans, peppers, tomatoes, strawberries and red cabbage). This facility is equipped with several sets of **sensors** that provide real time information related to soil and ambient temperature, humidity and air quality, using a purpose-built **IoT infrastructure** based on LPWAN networks. These data can be monitored remotely within the proposed digital environment, which can be accessed through a web browser. Additionally, **intelligent data analysis** allows for early detection of abnormal events and activation of warnings and alerts.

A **virtual agent** has been embedded in this setting, playing the support role of a guide to provide information and facilitate the decision-making process. During the design of this virtual persona, **socially relevant factors** were taken into account, as well as the realism of these social attributes (i.e. social fidelity) to emulate natural, real-world interaction with the end user. Functional requirements were defined, a customized, XR-ready animated avatar was created, and an understandable content flow was developed with the help of two agri-food experts, so as to deliver an innovative and helpful experience. Wording has been carefully designed for clarity, simplicity and ease of understanding for any user not familiar with these sort of solutions.

To provide natural means of communication, this virtual agent is supported by **Natural Language Processing technologies**, such as Automatic Speech Recognition to allow the user to command by using their voice, and Natural Language Generation to produce a spoken response. This integration of voice capabilities establishes a dynamic and accessible communication flow for any user, regardless of their digital skill. In addition, the virtual agent is enhanced with lip sync to increase lifelikeness, along with several state-dependent animations that render the interaction more natural while providing feedback to the user, **emulating non-verbal language** (i.e. waving while welcoming the user).

The underlying **conversational AI system** that provides this agent with a "brain" operates as a chatbot with several stages. The system attempts to classify the user's input with intent recognition in order to identify a question mapped to a predefined answer. When unsuccessful, it then generates extractive information out of the context the AI has been trained on. Generative AI models were also tested, but their adoption for this use case was dismissed due to an excess of *hallucinations*, as well as producing a number of inappropriate responses, making it harder to reach the desired information.

Results

An **IoT infrastructure based on LPWAN** technologies has been deployed, along with the sensors required, designed to monitor key crop related variables in real time. This provides the whole area with a **communications infrastructure suitable for a rural environment**, which can be scaled up by adding new data sources over time.

The **digital twin of the agrarian landscape of Peón** (rural area in Villaviciosa, Spain) provides a friendly and accessible way to explore the territory and its features by means of a visual and engaging web application. It is also populated with several points of interest that depict and monitor the **hyper-sensorized target facilities**.

Figure 72. Landscape and target facilities digitization (left), virtual/real offices building comparison (right).

The **CTIC RuralTech digital twin** represents our research and development offices specialized in rural innovation and smart territories. This building is equipped with a home automation system that monitors and controls the state of its dependencies. This way, room presence, lighting and temperature **data is integrated and spatially contextualized in the digital replica**, providing an easy to understand, centralized interface that allows for operation from within (i.e. lights can be handled remotely).

Figure 73. Office digital twin, home automation data monitoring and manipulation.

The **climate simulator digital twin** is also seamlessly integrated into the digital landscape, likewise embedding soil and ambient temperature, humidity and CO_2 data for remote monitoring. Critical levels for each of these parameters have been identified, so that real time warnings are triggered within the virtual environment in order to **alert the user in case of anomalies**.

Figure 74. *Climate simulator digital twin exterior (left), crop data visualization (right).*

The **virtual agent** is meant as an interface to ease the handling of the whole environment and guide the user within the application, providing **comprehensibility and transparency**. This character supports the user throughout the whole experience; being able to solve questions and provide detailed information **by means of natural language, both spoken and written.** Whether it is to unravel details about the area, the several points of interest, sensor data or general information about technologies involved, the character accompanies the user on an educational and immersive tour within this environment. Likewise, it can report any alerts detected to assist in troubleshooting, helping with their interpretation and notifying about critical events while also providing **recommendations for corrective actions**. This proactive approach allows users to understand the underlying logic of the warnings and to take informed actions accordingly, promoting active participation and greater control out of the virtual environment. A core feature of this agent is the **embodiment of a well and dearly known local personality**, Ramón Álvarez de Arriba, benefactor and donor to the region with many historical contributions to its development. By giving the character a friendly, approachable and recognizable appearance, connection and engagement with local agents is ensured, while social fidelity among users is increased.

Figure 75. *Ramón Álvarez de Arriba's picture (left), his embodiment as virtual agent (center), integration within the virtual environment application (right).*

Lastly, the **WebXR compliant deployment** of the virtual environment allows for convenient availability using any kind of device or platform through the web, thus **maximizing accessibility** and allowing the solution to potentially reach any user, at any location. It empowers users, both those digitally proficient and newcomers, to seamlessly access and leverage the benefits of this immersive experience across a diverse range of contexts. WebXR integration extends to devices that allow for Virtual Reality capabilities, featuring an **additional VR presentation and interaction mode**, underlining this solution's adaptability and commitment to making the experience accessible to all. By delivering several presentation modes suited to the capabilities of multiple sorts of devices, an optimal experience can be ensured for every user.

Figure 76. *Virtual Reality WebXR visualization of the environment on Meta Quest 2 browser.*

Conclusions

This paper presents a comprehensive solution that leverages WebXR standards to address the challenges faced by farmers in the context of AI-driven agricultural management. By integrating immersive technology and virtual agent guidance, this solution aims to promote trust, transparency, and comprehensibility, empowering farmers to embrace AI and enhance their agricultural practices under sustainability standards. This work contributes to the advancement of smart farming practices, benefiting both agricultural productivity and environmental stewardship.

Future Research Directions

Considering the focus of the present work, below are highlighted some relevant points, related to both user-centered design and technological development issues, as well as technical matters, to be taken into account when proceeding with future research.

Firstly, it would be interesting to assess the perceived level of social fidelity, aiming to align this aspect with user preferences and expectations, thereby enhancing trust, adoption, and use of such solutions. Active user engagement will be essential for this purpose.

In addition, it is relevant to evaluate the current level of technological acceptability of the final solution, especially concerning its transparency, in order to realize its inherent virtues and advantages.

Furthermore, given that the current solution relies on NLP, and considering recent developments in this field, exploring the use of Large Language Models linked with dialects existing in rural areas could be of interest to enhance usability and comprehensibility.

Lastly, a significant technical challenge towards promising research directions relates to the scalability and portability of this solution to other territories. Procedural generation of the terrain model based on GIS data will be considered in order to support dynamic regions of interest that display information from geolocated IoT sensors, ensuring that a solution as the one proposed could be applied elsewhere.

Acknowledgements

This research has been funded by the Spanish Ministry of Science and Innovation through the Centre for the Development of Industrial Technology (CDTI) under the project CEL.IA (CER-20211022), which aims to promote the application of Artificial Intelligence in different sectors such as Industry, Energy, Health, Transport and Territory. To achieve this goal, five Spanish RTOs work in a coordinated manner to facilitate the transfer of knowledge and innovative solutions.

MINISTERIO DE CIENCIA E INNOVACIÓN CDTI

@CDTIoficial

References

Alam, M. A., Ahad, A., Zafar, S., & Tripathi, G. 2020. A Neoteric Smart and Sustainable Farming Environment Incorporating Blockchain-Based Artificial Intelligence Approach. In Cryptocurrencies and Blockchain Technology Applications (pp. 197–213). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119621201.ch11

Haque, A. B., Islam, A. K. M. N., & Mikalef, P. 2023. Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. Technological Forecasting and Social Change, 186, 122120. https://doi.org/10.1016/j.techfore.2022.122120

Kyrlitsias, C., & Michael-Grigoriou, D. 2022. Social Interaction With Agents and Avatars in Immersive Virtual Environments: A Survey. Frontiers in Virtual Reality, 2. https://www.frontiersin.org/articles/10.3389/frvir.2021.786665

Sinatra, A. M., Pollard, K. A., Files, B. T., Oiknine, A. H., Ericson, M., & Khooshabeh, P. 2021. Social fidelity in virtual agents: Impacts on presence and learning. Computers in Human Behavior, 114, 106562. https://doi.org/10.1016/j.chb.2020.106562

Tiainen, M. 2021. To whom to explain and what? : Systematic literature review on empirical studies on Explainable Artificial Intelligence (XAI). https://www.utupub.fi/handle/10024/151554

WebXR Device API. (n.d.). Retrieved 10 August 2023, from https://www.w3.org/TR/webxr/

Weitz, K., Schiller, D., Schlagowski, R., Huber, T., & André, E. 2021. "Let me explain!": Exploring the potential of virtual agents in explainable AI interaction design. Journal on Multimodal User Interfaces, 15(2), 87–98. https://doi.org/10.1007/s12193-020-00332-0

Developing advanced automotive user interface using augmented reality

Hansung Lee, Byeongjun Choi, Ilwan Kim, Jinho Son, and Jongtae Park

Smart Mobility Laboratory, B2B Advanced Technology Center, Chief Technology Office, LG Electronics, South Korea

Corresponding author: hansung.lee@lge.com

Keywords: Augmented Reality, Automotive, User Interface

Introduction

Everyone must have had the experience of getting out of the wrong exit at one time or another because they couldn't find the correct exit at the roundabout. In the automotive industry, in order to provide useful information to drivers, a technology that combines real time video and navigation information and expresses them using augmented reality is being developed and mass-produced. In this paper, overall architecture and basic features of automotive augmented reality guidance will be introduced. Furthermore, by applying intuitive Pathfinder concept, advanced features of automotive augmented reality guidance UIs such as turn UI, roundabout UI, and lane level guidance UI will be explained.

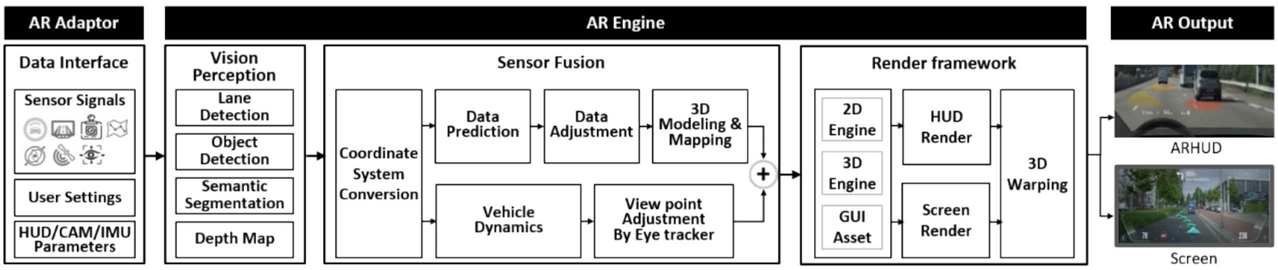


Figure 77. Architecture of automotive augmented reality guidance system

Overall architecture of automotive augmented reality guidance

The AR adaptor receives and parses various vehicle sensor signals, reads user settings and calibration parameters for HUD or camera, and delivers them to the AR engine by the AR protocols. The AR engine analyzes the sensor data delivered through the AR adaptor to recognize the surrounding environment, calculates the location to project the AR guidance UI, and renders it. The AR output rendered through the AR engine is finally displayed on several display screens such as HUD (Head Up Display), cluster, CID (Central Information Display), PD (Passenger Display), RSE (Rear Seat Entertainment), etc. The first figure shows the architecture of automotive augmented reality guidance system.

In order to place AR guidance UIs in the augmented world, the information of vehicles and surrounding objects must be converted into an AR coordinate system. The following figure shows the process of converting the location information of objects acquired from the ADAS and GPS coordinate system into the AR coordinate system (ar3d), a unified world coordinate space for AR, and finally converting it into the front view camera coordinate system (fv3d).

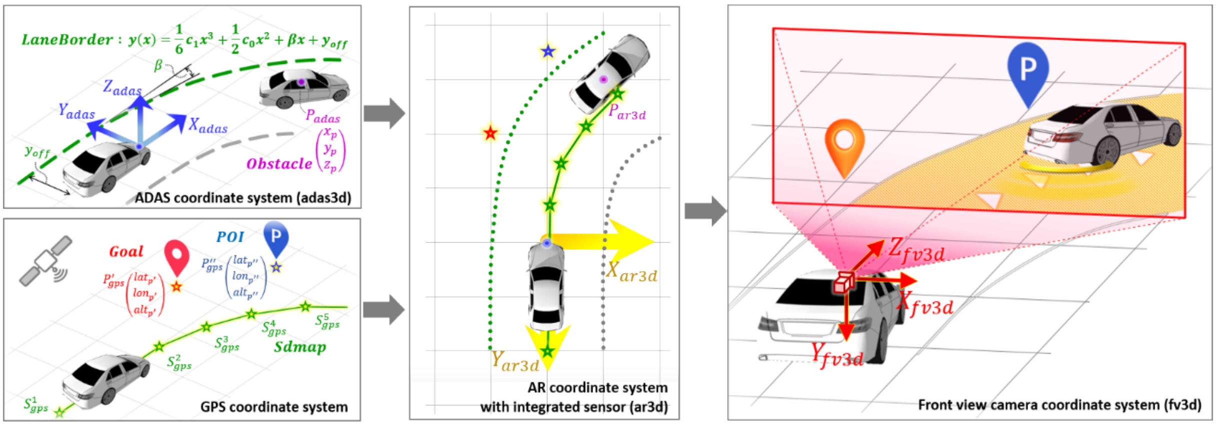


Figure 78. AR coordinate system conversion

Basic features of automotive augmented reality guidance

The basic features related to navigation in AR guidance are a carpet that highlights the lane the ego vehicle is driving on, a TBT (Turn By Turn) that is displayed above the driving lane, a goal that visualizes the destination location, and a POI (Point Of Interest) that shows major facilities such as restaurants, gas stations, and parking lots. The basic features related to ADAS in AR guidance are FCW (Forward Collision Warning), which visualizes a collision warning with a vehicle ahead, LDW (Lane Departure Warning), which visualizes a lane departure warning when the ego vehicle approaches the lane boundary, BSD (Blind Spot Detection), which visualizes a warning about obstacles in the ego vehicle's blind spot, and PD (Pedestrian Detection), which visualizes a collision warning with pedestrians. The following table shows basic features of AR guidance.

Table 6. Basic features of AR guidance

Carpet	TBT	Goal	POI
			
FCW	LDW	BSD	PD
			

Advanced features of automotive augmented reality guidance

Pathfinder concept: Pathfinder is a 3D UI composed of spade and chevron. Spade represents the vehicle condition (driving direction, speeding, etc.), and chevron represents the road guidance (maneuver point, driving trajectory, etc.). In the following content, we look at UIs for various route guidance using Pathfinder.

Figure 79. Pathfinder concept

Turn UI: In general, route guidance at intersection was simply displayed as static guidance of the remaining distance to the maneuver point and the direction to turn using the TBT UI. However, it is possible to help drivers find the exact position and direction to turn through more natural and intuitive chevron animation using Pathfinder. The following figure shows the process of the improved turn UI using Pathfinder.

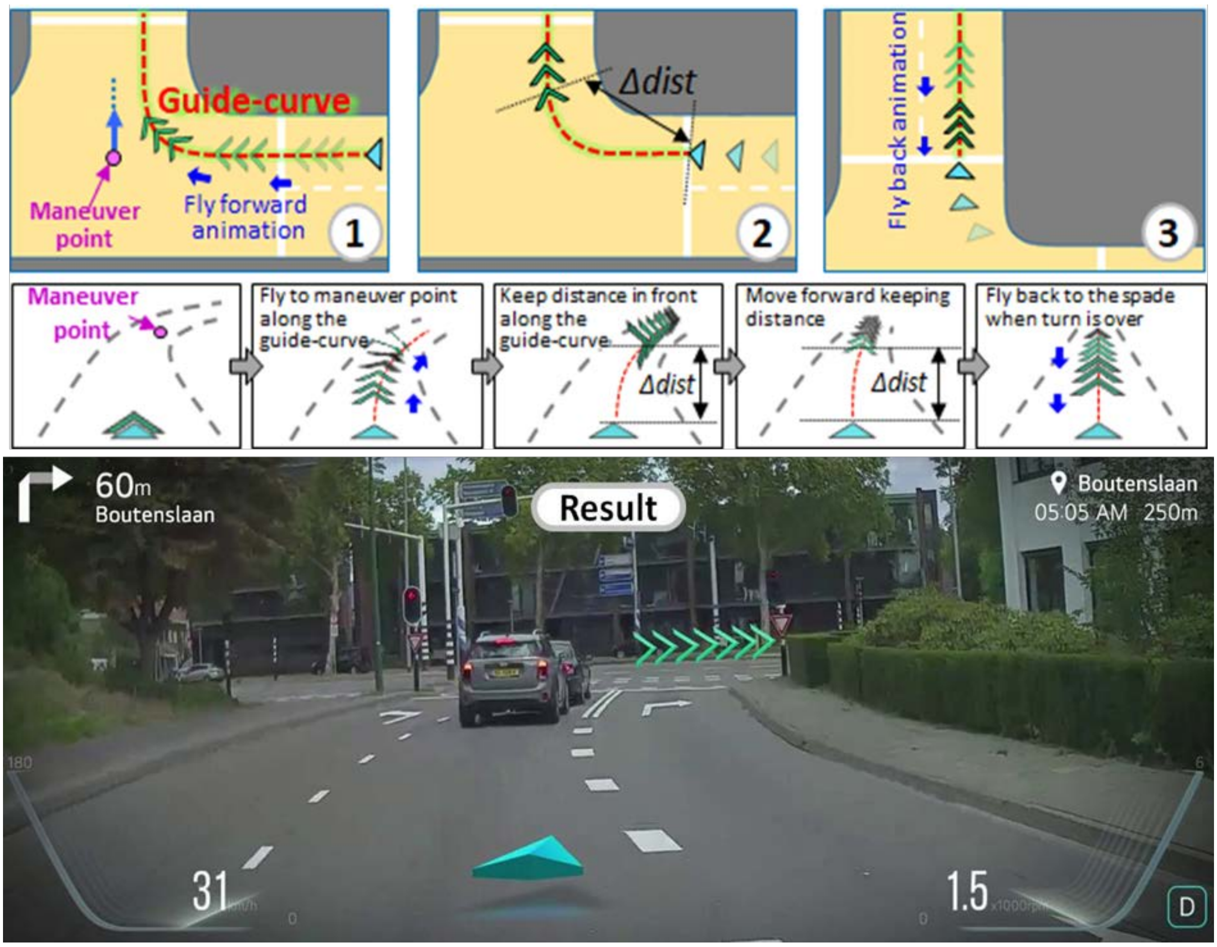


Figure 80. Turn UI using Pathfinder

Roundabout UI: Existing route guidance at roundabout was simply expressed through the TBT UI for entry and exit points in a similar way to intersection. The roundabout also uses Pathfinder to express the driving path from the entry point to the exit point of the roundabout through chevron animation, helping to find the way more easily. The following figure shows the process of improved roundabout UI using Pathfinder.

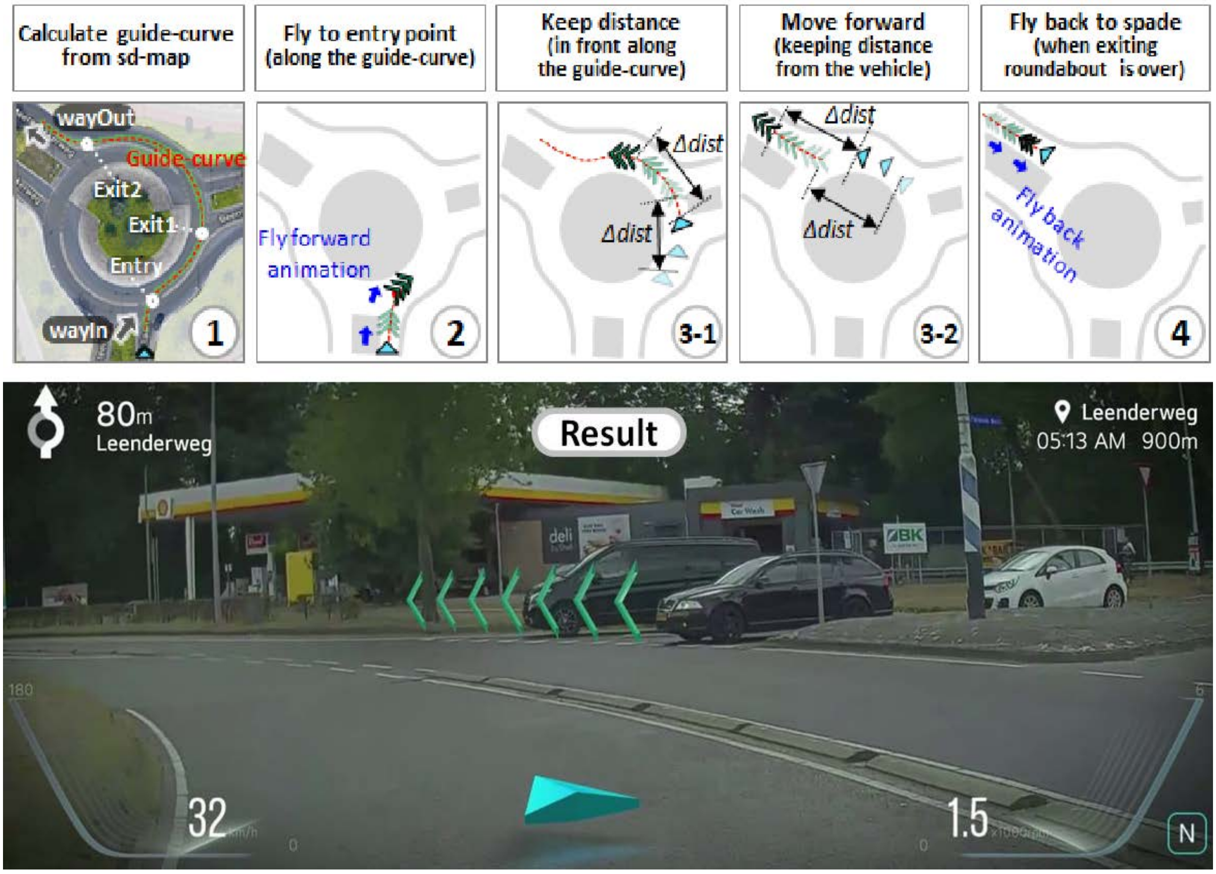


Figure 81. Roundabout UI using Pathfinder

Lane level guidance UI: Existing route guidance on highway was expressed only with the remaining distance to the exit and the exit direction through the TBT UI. In order to get out of the exit without missing the exit while driving on highway with many lanes, route guidance for changing lanes at the right time is required. For lane level guidance, it is essential to recognize the lane position of the ego vehicle and determine the timing of lane change guidance. The following figure shows the overall architecture for lane level localization and guidance. The lane border crossing detector detects whether the ego vehicle has crossed the lane border or not, and also recognize the direction of lane change and use this information for determining which lane the ego vehicle is located in. The lane detection information processor uses lane detection result from ADAS for providing clue for the location of the ego vehicle. The vehicle detection information processor takes the result of vehicle detection information from ADAS, and use it to identify surroundings of the ego vehicle. The lane level guidance uses information from the lane level localization, navigation, ego-vehicle, and ADAS to make decision whether to change lane or not. For example, since changing lanes in a solid lane marker section is illegal, lane change guidance is not provided in this section based on the lane detection from ADAS.

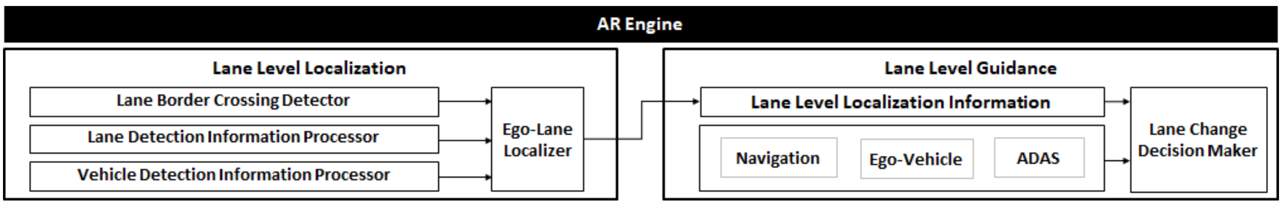


Figure 82. Architecture of lane level localization and guidance

The following figure shows Pathfinder UI for lane change and the experimental result of lane level localization and guidance on highway. The bottom left of the result screen indicates that there are one lane on the left of the ego vehicle and two lanes on the right of the ego vehicle, which means that the ego vehicle is located on the second lane from the left on the road with four lanes. Decision to change lane to the right has been made based on the remaining distance to the point of highway exit.

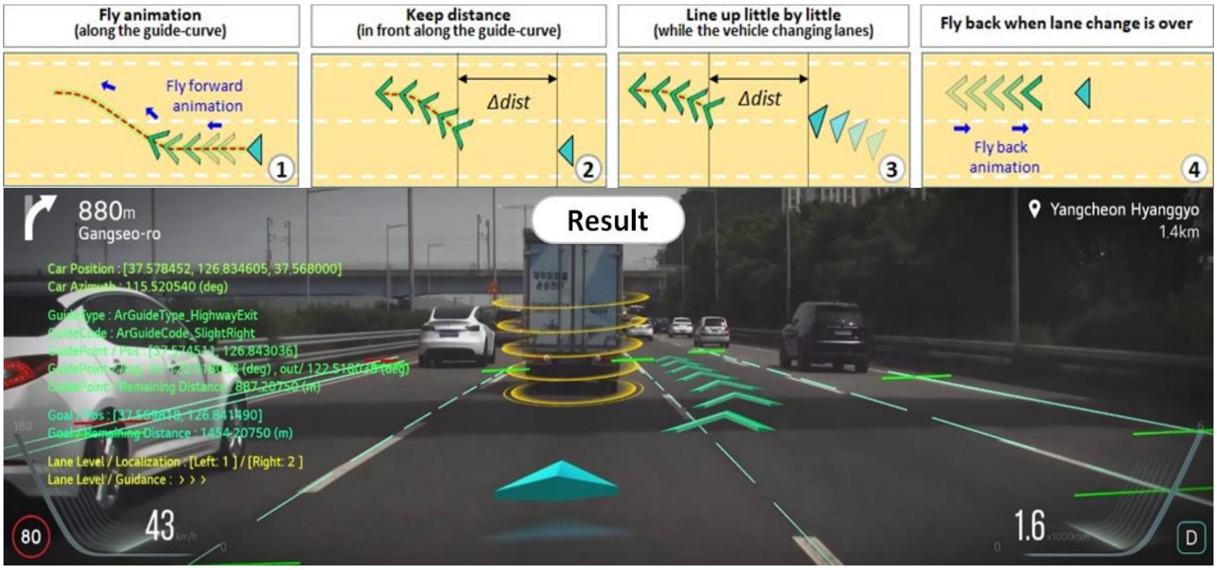


Figure 83. *Pathfinder UI for lane change and experimental result of lane level localization and guidance*

3D POI: Using a 3D rendering engine, we can display AR with more flashy and cool effects than the existing 2D rendering engine. To display 3D POIs, the GNSS coordinates of each 3D POI are converted into AR coordinates, and then the 3D object is rendered at the projected location on the front camera coordinate system. The following figure is the result of displaying POIs such as cafes and restaurants and air balloons for advertising in 3D AR.

Figure 84. *3D POIs and air balloon*

Removal of HMI overlap using semantic segmentation: When displaying AR UI such as route or lane guidance or POI by augmenting it on the front camera image, there are cases where major objects around are obscured by the AR UI. To this end, by using the semantic segmentation results obtained through vision perception, it is determined whether the AR UI overlaps with the main object, and then the AR overlap is removed through various methods such as moving the position of the AR UI or removing the overlapped part through smoothing and translucency effects. The figure below shows a case where the vehicle in front is obscured by an AR carpet, and is the result of the overlap removal method using smoothing and translucency effects.

Figure 85. Removal of AR carpet overlap using semantic segmentation (left: overlapped, middle: overlap removed, right: overlap removed by smoothing and translucency effects)

Conclusion

We proposed advanced automotive user interface that provides useful information to drivers using augmented reality. We designed the AR system architecture that analyzes various vehicle sensor data to locate the ego vehicle in its surroundings and draws AR guidance UIs. In addition, we have developed advanced features as well as basic features, and through this, we have been able to provide more natural and intuitive automotive user interface using augmented reality.

Future works

We plan to conduct real-vehicle driving validation for the lane level localization and guidance function on major courses in Europe and South Korea. Also, we will continue to develop more advanced and useful features such as Surround AR that provides AR guidance in various fields of view, integrated display of AR and 3D navigation, glass AR navigation, and HD map based lane level localization and guidance.

References

Antonov, S., Fehn, A., & Kugi, A. 2011. Unscented Kalman filter for vehicle state estimation. Vehicle System Dynamics, 49, 1497–1520. https://doi.org/10.1080/00423114.2010.527994

IEEE Intelligent Vehicles Symposium (IV), 1152-1157.

Rabe, J., Hübner, M., Necker, M. & Stiller, C. 2017. Ego-Lane Estimation for Downtown Lane-Level Navigation. https://doi.org/10.1109/IVS.2017.7995868

ProuVR: A Collaborative Design and Evaluation Virtual Reality Platform

Nawel Khenak[1], Yiran Zhang[1], Cédric Bach[1], and Guillaume Jégou[2]

[1]Human design Group, France,
[2]b<>com, Cesson-Sévigné France

Corresponding authors: nawel.khenak@humandesign.group, yiran.zhang@humandesign.group, cedric.bach@humandesign.group, Guillaume.jegou@b-com.com

Keywords: Operators Behavior, Human Factors, Production System Design, Virtual Reality, Data Analysis

Introduction

Abstract

In this paper, we introduce ProuVR, an immersive and collaborative platform aimed at optimizing design-evaluation cycles across diverse workplaces. ProuVR is based on b<>com *Ngagement* technology, which empowers ergonomists to simulate real-world scenarios within virtual environments, facilitating in-depth analysis of the behavior and interactions of future end-users within their respective work environments. A user test was conducted with 15 participants to assess the usability of the platform. The task involved an upgrade of a control room for operators and analyzing their behavior to validate the room design. Both quantitative and qualitative data were collected during the evaluation. The data analysis primarily focused on users' well-being and overall experience with ProuVR. The results highlighted that ProuVR maintained users' comfort during its utilization. Moreover, users highly valued the platform's novelty and functionalities, especially the 3D interactive tools it offers. However, the usability level of the platform was assessed as moderate, indicating the need for further improvements. As a result, recommendations have been extracted from the users' feedback, which will be carefully integrated into the next generation of ProuVR.

Introduction and Rational

Virtual reality (VR) is increasingly used to improve *ergonomic* practices in diverse fields such as education, aviation, automotive, assembly, and maintenance (Brown et al., 2021; da Silva et al., 2020). VR tools can simulate high-fidelity virtual environments (VEs) in which observing participants are fully immersed and able to interact with a synthetic workplace, in a controlled manner, without requiring direct interaction with the actual environment (Makransky et al., 2019; Chavan, 2016). Ergonomists have long been interested in using such VEs in ergonomic *design and analysis* (Wilson, 1997). Indeed, many of the ergonomic limitations that are associated with traditional design approaches such as determining optimal placements for displays and controls, as well as identifying suitable positions for sightlines and illumination, can be effectively simulated and modeled using VR (Foster and Burton, 2004). In addition, VR tools can provide support for the analysis of ergonomic measurements of workplaces including workload, situational awareness, and information processing. Then, these tools have the potential to enhance the development of novel processes, such as maintenance procedures in the aviation industry (Brown et al., 2021). Moreover, if correctly tailored to the context of use, VR allows for optimizing costs and productivity without affecting safety (la Monaca et al., 2019; Vance, 2017).

The advantages that VR may offer have led to the development of "**ProuVR**", an **immersive and collaborative platform** for optimizing **user-centered design (UCD)** cycles across diverse workplaces (Devreux et al., 2023). ProuVR offers Human Factors solutions based on b<>com *Ngagement* technology (Calvet et al., 2020), which is the result of a research and development collaboration between Human design Group (hdg) and Technology and Research Institute b<>com. The project aims to leverage the flexibility of VR to support UCD. Precisely, ProuVR enables ergonomists to simulate operational scenarios based on the future working environment and conduct detailed human factors analyses to improve/validate the design.

In this paper, we present a technical overview of the ProuVR platform, followed by a user test aimed at evaluating its usability. The task performed in the test actively involves participants in both the design and evaluation phases and has been customized to align with the industrial needs. The main purpose of this test is to collect feedback to improve the maturity of the platform.

Platform Overview

ProuVR consists of two main modules:

1. A cross-device **immersive multi-user environment** (see Figure 86 left) for efficient 3D workplace design. Users can use PCs or VR systems to work collaboratively in a shared 3D environment, through:
   - Design tools (e.g., rangefinder, accessibility template, mannequins, pedestrian flow simulator, embedded 2D web pages, etc.).
   - Advanced interaction features (e.g., teleportation, object manipulation and modification, save and load environments, recording notes, etc.).
   - A library of customizable 3D objects (e.g., offices, cockpits, dashboards, screens, chairs, etc.).
2. and a **2D interface** (see Figure 86 right) associated with a **BITalino**[13] **toolkit** for monitoring VR users, through:
   - Capturing behavioral data (e.g., areas of interest) and physiological data (e.g., heart rate, electrodermal activity, and pupil size) in a minimally intrusive way, without interrupting the task at hand.
   - Continuous analysis of a user's cognitive state based on data captured.
   - Real-time visualization and recording of users' views.
   - Collecting and exporting data, and replaying recorded views to support post-analysis.

[13] https://www.pluxbiosignals.com/collections/bitalino

Figure 86. (Left) An example of a 3D simulated control room. (Right) The monitoring interface with different users' views and physiological data.

Usability test and results

To assess the usability of ProuVR, a user test was conducted with a focus on **a control room design task**. The aim was to simulate a representative use case and evaluate the platform's usability in an operational setting. The task was divided into two phases: (1) **designing an upgraded control room** that met predefined requirements and (2) **verifying its ability to match the operator's needs**.

Experiment setup

The VR setup consisted of HTC Vive Pro Eye with position and orientation tracking in a 2.5m x 2.5m area. The VE was rendered using Unity Game Engine. In addition, an audio communication system (i.e., Microsoft Teams) was set up to support remote collaboration.

Participants

The user test involved 15 participants (8 men and 7 women), aged between 21 and 41 (mean = 29.6, *std* = 5.58). All the participants had a background in ergonomics but were novices to the ProuVR platform. Twelve out of the 15 participants had prior VR experience, while six of them reported experiencing discomfort when using VR devices.

Experimental task

In the *design phase*, participants who assumed the role of "*ergonomists*" were assigned **the task of designing an upgrade of a control room by integrating four additional workstations**. They were given several criteria to guide their design process. These criteria included ensuring the ability to share information between two operators and facilitating operator circulation within the control room. Moreover, the design should also meet ISO standards, such as ensuring that the worktop height falls within a specific range. During this phase, participants were invited to work in pairs in a shared virtual control room. One participant operated from a PC, while the other used the immersive system. Participants had access to ProuVR design tools, allowing them to add/remove objects, resize/reposition them, and utilize tools such as the rangefinder to assess the design from a human factor's perspective.

In the *evaluation phase*, a third participant was added, taking on the role of the "*operator*". The main objective was to evaluate the newly upgraded control room from an end-user perspective in VR immersion. This involved **assessing the design with a focus on the operator's experience** and considering requirements that might have been overlooked during the earlier stages. This third participant was thus given additional criteria that were different from those provided to the participants in the *design phase*. As an example, one requirement given to the third participant was to ensure that the design of the control room does not obstruct the visibility of the screens including the periphery of the control room. During this phase, one of the two ergonomists from the *design phase* accompanied the operator to evaluate the virtual control room using VR equipment. Simultaneously, the other ergonomist monitored their activities on a PC and tracked the operator's physiological data and cognitive states in real time.

Procedure and Data Collection

Upon arrival at the sites, participants signed an informed consent form and filled out a demographic questionnaire. For the well-being of the participants, a cybersickness briefing was performed. Participants were then given 20 minutes to configure the platform using a provided user manual, followed by an additional 20 minutes of training. After a 10-minute break, participants received task instructions and VR users completed a first **Simulator Sickness Questionnaire (SSQ)**. To ensure participant well-being, the VR exposure time for the task was limited to a maximum of 23 minutes in a row based on previous studies. After completing the task, participants filled out a second SSQ to evaluate changes in cybersickness. They also completed the **System Usability Scale (SUS)** and **User Experience Questionnaire (UEQ)** to assess platform usability and user experience. Finally, participants were provided with an **open-ended questionnaire to list three platform advantages and disadvantages**. Each phase of the experiment lasted approximately two hours.

We registered 5 trials for each phase. In each phase, we measured the **task completion time (TCT)**, the **number of times participants requested assistance**, and the **number of errors made** while attempting to achieve a goal.

Results

In the *design phase*, the TCT was found to have **a mean of 22.60 minutes** (*std = 3.91*). On average, **participants required assistance 8.9 times** (*std = 6.32*) and **made 9 errors** (*std = 6.16*). **No significant difference in cybersickness** levels was found before (*mean = 15.71*, *std = 17.42*) and after (*mean = 15.71*, *std = 20.55*) being immersed. The usability level, as evaluated using **SUS, was found to be medium** (*mean = 56.51*, *std = 14.96*). According to the UEQ, the platform received **good to excellent scores in the dimensions of novelty** (*score = 1.30*) and **stimulation** (*score = 1.78*). However, the participants reported **lower scores for perspicuity** (*score = 0.18*) and **dependability** (*score = 0.55*) dimensions. Thematic analysis of user feedback revealed the presence of 14 codes, organized into four themes: *system capacity*, *VR content*, *interaction features*, and *user motivation*. Participants particularly emphasized the **realism of the 3D virtual content** and the **efficiency of the collaboration** facilitated by the system. However, they also reported **difficulties when using the interactive tools provided and manipulating objects**, especially when the objects were positioned close to each other.

In the *evaluation phase*, participants completed the task in **an average time of 20.77 minutes** (*std = 4.30*), **requested assistance 10.5 times** (*std = 7.17*), and **made 4.34 errors** (*std = 4.71*). However, participants reported **experiencing more symptoms of cybersickness** after engaging in the immersive experience. The mean cybersickness score increased to 27.78 (*std = 41.13*) compared to the baseline measurement where the mean score was 17.10 (*std = 15.40*). The **usability of the system was found to be medium** (*mean = 58.17*, *std = 14.51*). The platform received **positive scores in the dimensions of stimulation** (*score = 1.15*) and **novelty** (*score = 0.85*), but **lower scores in the dimensions of attractiveness** (*score = 0.63*) and **dependability** (*score = 0.53*). Thematic analysis of user feedback led to the identification of 17 codes organized into five distinct themes: *system capability*, *VR content*, *3D interaction features*, *2D interaction features*, and *user motivation*. While participants **highly appreciated the realism and efficiency** that ProuVR provided, they also mentioned the **instability of the system**.

Conclusion and future work

In this paper, we presented "*ProuVR*", an immersive and collaborative platform to optimize UCD cycles across diverse simulated environments, based on b<>com *Ngagement* technology. A user test of the platform inspired by market needs was conducted. Results highlighted that ProuVR maintained users' well-being during the design phase, and less during the analysis phase even if the cybersickness score remains associated with moderate symptoms (Stanney et al., 2020). Moreover, users highly valued the platform's novelty and functionalities, especially the 3D interactive tools it offers. The flexibility provided by the VR platform to create immersive workplaces was also appreciated by users. However, the usability of the platform, assessed as moderately satisfactory based on Bangor's scale (Bangor et al., 2009), indicates room for further improvement. Therefore, recommendations were extracted from the users' feedback related to system capability (e.g., integrating an effective way of alerting users when there is a risk of actual collision), VR content (e.g., implementing audio communication between immersed users), and 3D/2D interaction features (e.g., adding different units of measurement and incorporating shortcuts). Considering feedback from the test will allow specifying new interactive tools required for co-design and analysis as well as implementing new solutions for preserving user comfort and ensuring a good user experience of the platform.

Further tests will focus on investigating how ProuVR can support the conception and analysis of ergonomics in more specific tasks, such as assembly and training tasks. In particular, the tests will have to evaluate the benefits of using the VR platform to increase user performance.

References

Bangor, A., Kortum, P., & Miller, J. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. Journal of usability studies, 4(3), 114-123. https://uxpajournal.org/wp-content/uploads/sites/7/pdf/JUS_Bangor_May2009.pdf

Brown, C., Hicks, J., Rinaudo, C. H., & Burch, R. 2021. The use of augmented reality and virtual reality in ergonomic applications for education, aviation, and maintenance. Ergonomics in Design, https://doi.org/10.1177/10648046211003469.

Calvet G., Jégou G., & Lewkowicz D. 2020. Egagement, une plateforme de design et d'évaluation rapide. In Proceedings of 55ème congrès SELF.

Chavan, S. R. 2016. Augmented reality vs. virtual reality: Differences and similarities. Int. J. Adv. Res. Comp. Eng. Technol. 5(6), 1947–1952

da Silva, A. G., Winkler, I., Gomes, M. M., & Pinto, U. D. M. 2020. Ergonomic analysis supported by virtual reality: a systematic literature review. Proc 22nd Symposium on Virtual and Augmented Reality (SVR) (pp. 463-468). IEEE. https://doi.org/10.1109/SVR51698.2020.00074

Devreux G., Lahoual D., Zhang Y., Calvet G., Jegou G., & Gody A. (Under publication). Retours d'expériences de la plateforme PROUVR. Ergo'IA 2023 conference.

Foster, P., & Burton, A. 2004. Virtual reality in improving mining ergonomics. Journal of the Southern African Institute of Mining and Metallurgy, 104(2), 129-133. https://hdl.handle.net/10520/AJA0038223X_2873

la Monaca, U., Bertagna, S., Marinò, A., & Bucci, V. 2020. Integrated ship design: An innovative methodological approach enabled by new generation computer tools. International Journal on Interactive Design and Manufacturing (IJIDeM), 14(1), 59-76. https://doi.org/10.1007/s12008-019-00612-4

Makransky, G., Terkildsen, T. S., & Mayer, R. E. 2019. Adding immersive virtual reality to a science lab simulation causes more presence but less learning. Learning and Instruction, 60, 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007

Stanney, K., Lawson, B. D., Rokers, B., Dennison, M., Fidopiastis, C., Stoffregen, T., ... & Fulvio, J. M. 2020. Identifying causes of and solutions for cybersickness in immersive technology: reformulation of a research and development agenda. International Journal of Human–Computer Interaction, 36(19), 1783-1803. https://doi.org/10.1080/10447318.2020.1828535

Vance, J. M. 2017. Industry use of virtual reality in product design and manufacturing: A survey. Virtual Reality, 21(1), 1–17. https://doi.org/ 10.1007/s10055-016-0293-9

Wilson, J. R. 1997. Virtual environments and ergonomics: needs and opportunities. Ergonomics, 40(10), 1057-1077. https://doi.org/10.1080/001401397187603

AR coupled with localized simulation technology for Cultural Heritage

Georgios Karafotias, Tina Katika, Spyridon Nektarios Bolierakis, Giannis Karaseitanidis, and Angelos Amditis

Institute of Communication and Computer Systems (ICCS), Athens, Greece

Corresponding authors: giorgos.karafotias@iccs.gr

Keywords: Mobile AR, Augmented Reality, Object Tracking, Object Detection, Extended Tracking, Digital Heritage

Introduction

Cultural Heritage sites have started capitalising on the possibilities of AR and improving their visitor experience as more and more applications influence users' field applications (Aliprantis & Caridakin, 2019, Katika et al., 2022). The use of AR technologies in a real case scenario has been investigated by Pierdicca et al., 2016, in the Cardeto archaeological park in Italy, exploiting two different tracking systems: location-based and edge-based. Speed and response were recognized as the advantages of edge tracking, yet, their model relies on wireless sensor network which often is not available outdoors. Considering that large-scale archaeological sites face unique challenges compared to museums and exhibition areas Dragoni et al., 2019, used SLAM technology to superimpose virtual reconstructions at real scale. The lighting and user confidence with AR highly influenced their results to achieve matching between existing evidences and 3D reconstructions. In addition, their application depended heavily on data exchange and the lack of wireless connection constituted a major drawback. Finally, SLAM technology is energy consuming for the used devices.

Problem statement

Archeological sites are typically locations that their original view and appearance have undergone major changes and they pose certain characteristics, such as, the demand of large-scale 3D reconstruction, lack of internet connection, incapability of performing actions within the surroundings (for instance to install sensors), variations in lighting and textures (flora changes a lot), visitors of these sites have smartphone devices with low processing power, have very limited existing information to obtain via edge tracking techniques.

Research statement

An AR mobile app (i.e. CirculAR) has been developed focusing on state-of-the-art localization technology, aiming to address bottlenecks that previous efforts faced, such as large-scale 3D artefacts representation utilizing low energy and power from smartphone devices and operating offline. 3D geometry-based modelling was used for the design of the 3D artefacts representing the buildings located in the ruins of two archeological sites and one museum. More precisely, CirculAR aspires to support the visualization of 3D large-scale content, including information (text, audio, and video) in an accessible and inclusive manner; to use localization technology to reposition the 3D artifacts at their original position as accurately and as stable as possible; to seamlessly localize the 3D artefacts at field conditions with varying lighting, textures, possible absence of internet connection; to perform at common smartphone devices that have low performance in terms of energy and power.

CirculAR solution

CirculAR aims to provide a gamified experience enhanced with AR content in real-world locations, primarily targeted to Cultural Heritage sites. Curators, archaeologists, and other content creators can use a Content Management System (CMS) to create campaigns consisting of multiple geolocated points of interest (PoIs) with attached related multimedia files (text, images, videos, 3D models). CirculAR's users get a list of available campaigns to experience digital heritage in AR. After selecting an AR campaign with the desired content, the app switches to the map view where there are icons for the positions of the user and all the nearby POIs. If a PoI is selected, then information about it is displayed on the screen along with a button that opens the camera and a prompt cueing the user to start looking for a specific image, QR code or object.

The app is equipped with image and object trackers that recognize targets that are either precompiled or added in the runtime. In most cases, a characteristic object or one that is paramount to the scene is picked. The user is prompted to point the camera to the POI's appropriate target to initiate AR mode. Upon successful tracking of the specified target, the PoI's corresponding 3D model is spawned at the target's real-world location. The model is anchored to this position, so the user does not have to keep the target inside the camera's field of view. This type of tracking, called extended, is useful in most of CirculAR's use cases, since the digital content may be large enough to cover entire parts of archaeological sites (often tens of meters in size). As soon as the digital artefact is spawned, the user is able to observe it and, simultaneously, physically move relatively to the 3D model which is geolocated at the target's position in the tracking/initialisation phase. The mesh from the CMS is usually rendered in real-world scale so it matches the physical surroundings.

The application is equipped with three main views for the end-user:

1. The main AR mode as described above, where the user views the world augmented with the 3D content (Figure 87 and Figure 88). To enhance the app's interactability, the model may enable some of its parts to be tapped to trigger some kind of effects, such as animations. Furthermore, the model may be rotated with appropriate gestures. In the app screen, a series of buttons enable the end-user to: read information and instructions, toggle the narrational Text-To-Speech system for audio guidance, start a quiz associated with a PoI to assess their knowledge and understanding, reset the 3D model in case the anchoring has degraded or has accumulated calculation errors, and exit to the map view mode.
2. The view via a 3D viewer, where the PoI's model is rendered in the centre of the screen, and the user can use swipe gestures to rotate it and view it from any angle (Figure 89). The mesh has first gone through a sequence of model management calculations (colliders and bounding boxes) to assure it is scaled down so as to appear in its entirety.
3. A mobile VR mode that imitates a first-person camera where its rotational three degrees of freedom are controlled by the camera's gyroscope, while its translational three degrees of freedom are handled by corresponding UI buttons (forward, backward, left, right, up, down).

This mode allows the user to examine and walk through the real-scale model without having to physically move (Figure 90); an extremely helpful functionality in archaeological sites where some locations are not traversable or they extend beyond the limits of allowed visits.

All functionalities have been tested against a set of pre-defined key performance indicators (KPI) and their associated metrics by a team of beta testers (12 participants) at the Athens Epigraphic Museum. An example of the KPIs was to check if the user was able to interact successfully with the 3D augmentation model. The testers were able to assess the functionalities of object tracking, extended tracking, etc. At the same time, the testers successfully finalized the AR flow consisting of: profile editing, tutorial and demo experience, selection of campaign, object spawning, manipulation, rotation, scaling and viewing in 3 different modes, as well as, the gamification functionalities, including quizzes, scoring and rewards.

Future work includes a testing phase with a broader group of testers comprising of archaeologists, expert scientists and museum curators who already work in specific heritage sites and can provide accurate and timely feedback. The evaluation techniques (surveys, questionnaires, interviews, focus groups, and A/B testing) will be focused primarily on user experience. Moreover, the aim is to simultaneously assess the tracking algorithm's accuracy under diverse weather and lighting conditions as well as the stability of the rendered digital content when the trackable target object leaves (partially or fully) the camera's field of view.

Figure 87. *Viewing the 3D model in real scale.*

Figure 88. *Walking inside the 3D model.*

Figure 89. *Model in the 3D viewer.*

Figure 90. *Mobile VR mode.*

Acknowledgement

The research has been funded by APSIM project, which has received funding from Research – Create – Innovate program under T2EΔK-01745. The opinions expressed herein are those of the authors and the European Commission is not responsible for any use that may be made of the information it contains.

References

Aliprantis, J., & Caridakis, G. (2019). A survey of augmented reality applications in cultural heritage. Int. J. Computat.l Meth. Heritage Sci (IJCMHS), 3(2), 118-147. https://doi.org/10.4018/IJCMHS.2019070107

Dragoni, A. F., Quattrini, R., Sernani, P. & Ruggeri, L. 2019. Real Scale Augmented Reality. A Novel Paradigm for Archaeological Heritage Fruition. 1st Int. Interdisc. Conf. Digit. Environ. Educ., Arts and Herit., 657-670. https://doi.org/10.1007/978-3-030-12240-9_68

Katika, T., Karaseitanidis, I., Tsiakou, D., Makropoulos, C. & Amditis, A. 2022. Augmented Reality (AR) Supporting Citizen Engagement in Circular Economy. Circ. Econ. Sust. 2, 1077-1104. https://doi.org/10.1007/s43615-021-00137-7

Pierdicca, R., Frontoni, E., Zingaretti, P., Malinverni, E., Galli, A., Marcheggiani, E. & Costa, C. 2016. Cyberarchaeology: Improved Way Findings for Archaeological Parks Through Mobile Augmented Reality. International Conference on Augmented Reality, Virtual Reality and Computer Graphics, 172-185. https://doi.org/10.1007/978-3-319-40651-0_14

Plug and Learn: eXpeRtise at a Distance

Frank Ansorge[1], Elias Meltzer[1], Simon N.B. Gunkel[3], Rui Li[2], Fangtian Deng[2], Wenfeng Zhu[1], Rulu Liao[1], Bao Trung Duong[1], and Amelie Hagelauer[1]

[1]*Fraunhofer Institute for Electronic Microsystems and Solid State Technologies EMFT, Munich, Germany*
[2]*Fraunhofer Institute for Casting, Composite and Process Technology IGCV, Augsburg, Germany*
[3]*TNO, Networks, The Hague, The Netherlands*

Corresponding authors: frank.ansorge@emft.fraunhofer.de, elias.meltzer@emft.fraunhofer.de

Keywords: AR, Training and Education at a Distance, 3D Representation, Motion Tracking, Sensor Monitoring

Introduction

Industry worldwide is increasingly confronted with the challenge of highly qualified personnel leaving the company, and with them essential know-how and skills. This is especially true in the field of soldering. Particularly for aerospace applications, special skills in hand soldering are required so that the reliability of electronic components can be always guaranteed. Trained personnel are therefore of elementary importance to maintain the quality of production.

It is compelling to enable expertise or skills to be provided at a distance so that people can access high-quality training completely independent of location. Costly travel to training locations is eliminated, staff downtime can be significantly reduced, and training can be conducted at the same level at different production sites around the world. In addition to the industrial context (training, collaboration, inspection, maintenance), online education is key to engaging citizens with access barriers. However, the lack of interaction between trainers and trainees decreases learning success and readiness. The learning outcome and the willingness to learn decreases due to the lack of interaction between trainer and trainees.

Extended Reality (XR) applications are increasingly being integrated into training operations (Dede & Barab, 2009, Elfeky et al., 2021, Sampaio & Almeida, 2013). Often, these are (pre-recorded) specialized training courses. While it brings some benefits, there is still a clear need for more interaction here, according to Doolani et al., 2020. Novel immersive 3D platforms may enable new forms of remote education and training, but not all training procedures can be simulated with the high level of accuracy and sensory immersion required. For example, in ESA-certified high-end soldering training, students must feel the heat and judge the soldering fumes during training, while the instructor must evaluate the quality of the soldering result.

In this paper we introduce our immersive (AR enhanced) remote learning platform Plug and Learn with the specific focus on the use case of teaching hand soldering remotely to professionals. Further we outline four specific features that we believe have an important impact and improve the learning experience of students at a distance: photorealistic 3D representation of the teacher for natural interaction, trajectory comparison and augmentation of sensor information "on demand".

Plug and Learn – eXpeRtise at a Distance Platform

Mobile Learning Hub for Soldering

The basis of our eXpeRtise at a distance online training platform is an existing mobile soldering station (Mobile Learning Hub[14]). The station offers a camera system and a communication interface (2D using Microsoft Teams) in addition to a complete set of soldering-specific equipment according to Figure 91. Several soldering stations can be in operation at the same time, while the trainer can access the individual camera system and thus supervise and, if necessary, monitor each participant individually. The mobile soldering station allows for a full teaching experience resulting in the future into an ESA (European Space Agency), IPC or AVLE (Training association soldering technology electronics) compliant certification.

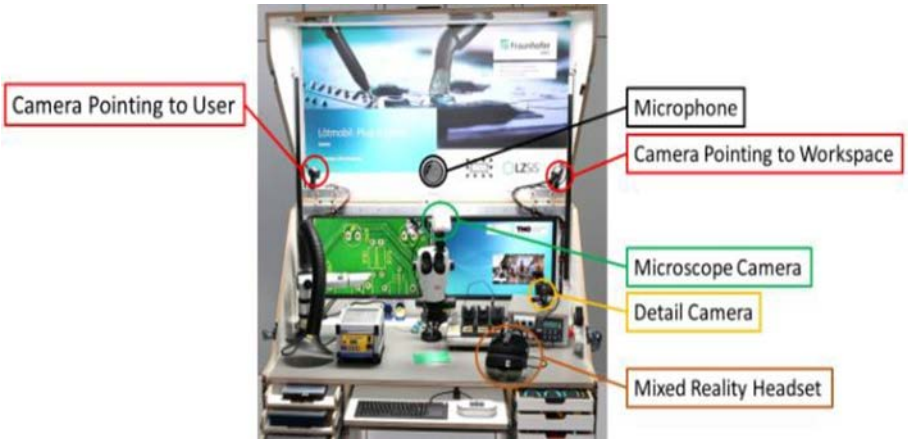


Figure 91. *Mobile Soldering station enhanced with XR technology. Black circle: microphone; Red circles: cameras pointing at user and workspace; Green circle: microscope; Yellow circle: detail camera; Brown circle: mixed realty headset (Microsoft HoloLens)*

3D representation

For remote teaching it is important that the individual communication between instructor and student is natural and engaging. For this we augment the view of a student with a photo- realistic 3D representation of the instructor into an immersive AR experience (Gunkel et. al, 2023). This is done on demand, e.g., only when a student wants to communicate directly with the instructor during practical exercises. This creates a sense of co-presence so that more natural as well as personal communication can take place. Figure 92 shows the setup of the *Plug and Learn* prototype. The final setup for the student consists of a PC running a Social XR web client to record the webcam feed from the soldering station, and a HoloLens 2 also running a Social XR web client to record the view from the HoloLens 2 into the system and to render the instructor's audio and the 3D user view. The trainer has the option to view multiple cameras that are equipped in the students Plug and Learn soldering mobile station. The different camera streams are displayed on the trainer and can be remotely operated and switched. A key aspect of the development is the simple and intuitive handling of the various AR functions. In this way, allowing flexible and individual student support, increased motivation, and the overall reduced need of support by the teacher.

[14] https://www.emft.fraunhofer.de/en/projects-fraunhofer-emft/remote-soldering-training-plug-and-learn.html

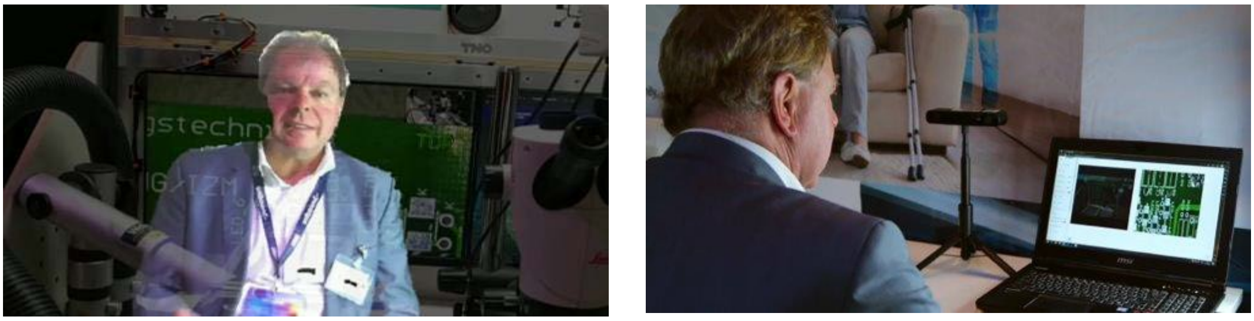


Figure 92. *Mobile Soldering station projection of the trainer in the student's field of view (left), Trainer in front of the camera at the trainer's desk (right), the camera feed is displayed on the monitor and can be switched remotely by the trainer (Gunkel et. al, 2023)*

Trajectory monitoring by optical motion tracking

Trajectory monitoring is achieved using Optical Motion Tracking Technology (OMT). The OMT solution comprises two core components that synergize: the optical camera and navigation passive markers (see Figure 93). The optical camera employs infrared (IR) light to precisely triangulate and determine the real-time pose of instruments (specifically markers) in a 3D space. The pose represents the position and orientation of the instruments with timestamp, which can be derived to a velocity and acceleration dataset. Its application spans diverse fields, including sports analysis, animation, virtual reality, robotics, and biomechanics, among others (Guerra-Filho, 2005).

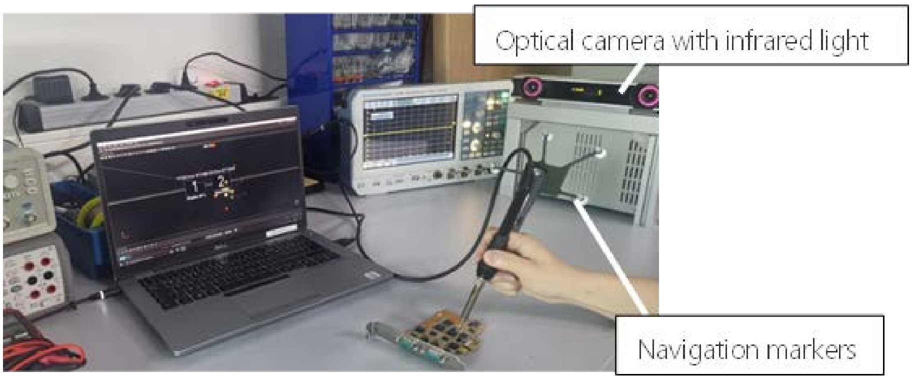


Figure 93. *Motion capturing of hand soldering by marker provided soldering iron and optical camera.*

In our scenario, instructors can record a soldering trajectory using a dummy soldering pen, which can later be visualized in a 3D space using Augmented Reality (AR) technology (Ribo et. al, 2001) to enhance student comprehension. Simultaneously, students can transmit real-time trajectories of their soldering pens to instructors. This facilitates prompt instructor feedback. Through comparing and data analysing the soldering trajectories of both instructors and students, it is possible to intuitive and digitally visualize whether students have mastered this skill. This real-time communication coupled with trajectory monitoring in 3D space significantly enhances the quality of remote training.

Sensor system for ambient monitoring

Additional information such as the ambient conditions during soldering (fumes, temperature, humidity) is measured by a sensor system. For this purpose, the sensor nodes are attached at various points on the mobile soldering station (extraction station, work area, etc.) in such a way that comprehensive monitoring is possible. The sensor data is pre- processed and, in the event of a critical condition, displayed in the student's field of view so that action can be taken. In this way, on the one hand, the conditions at the workplace can be made transparent and the safety of the participants can always be ensured remotely.

Conclusion and Future work

AR Technology is a powerful tool to enhance the motivation and learning success of students and professionals. Further, using AR technology, practical training can take place almost anywhere in the world - without time-consuming travel. Important information superimposed on real events creates significant improvement in the procedure of training compared to classroom training, where each student must be corrected for learned task individually. For example, marking objects as shown in Figure 94 could significantly improve communication when needed. In this paper we presented our immersive AR enhanced remote learning platform *Plug and Learn* with some key features that we believe will significantly improve the learning experience of our students. While getting all technology components together to allow the presented features in our platform we are currently planning dedicated expert session to shape the requirements of our solution further. Once ready we plan to test all features together with teachers and students to understand the actual impact of our augmented and immersive learning features.

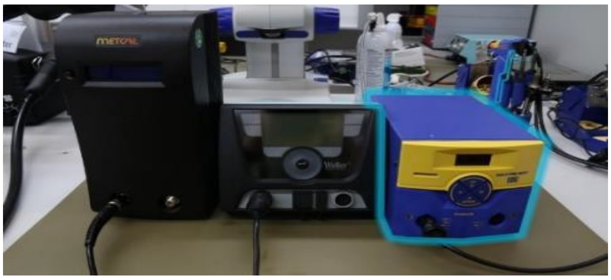


Figure 94. Easy Selection of Correct Solder Tool by Highlighting in Augmented Reality Glasses (Ansorge et. al., 2023)

References

Ansorge, F., Meltzer, E., Duong, B.T., Richter, L., Paul, G., Schröder, D., … & Kutter, C. 2023. Practical Training and Certification from a Distance - Enhanced by Augmented Reality. Festival der Zukunft, Deutsches Museum

Dede, C. & Barab, S. 2009. Emerging technologies for learning science: A time of rapid advances. J. Sci. Educ. Technol., 18, 301-304. https://doi.org/10.1007/s10956-009-9172-4

Doolani, S., Wessels, C., Kanal, V., Sevastopoulos, C., Jaiswal, A., Nambiappan, H., & Makedon, F. 2020. A review of extended reality (xr) technologies for manufacturing training. Technologies, 8(4), 77. https://doi.org/10.3390/technologies8040077

Elfeky, A. I. M. & Elbyaly, M. Y. H. 2021. Developing skills of fashion design by augmented reality technology in higher education. Interact. Learning Environ. 29(1), 17-32. https://doi.org/10.1080/10494820.2018.1558259

Guerra-Filho, G. 2005. Optical Motion Capture: Theory and Implementation. RITA, 12, 61-90.

Gunkel, S. N., Dijkstra-Soudarissanane, S., & Niamut, O. 2023. "You AR'right in front of me": RGBD-based capture and rendering for remote training. Proc. 14th Conf. ACM Multimedia Syst., 307-311. https://doi.org/10.1145/3587819.3593936

Ribo, M., Pinz, A., & Fuhrmann, A. 2001. A new Optical Tracking System for Virtual and Augmented Reality Applications. Conference Record - IEEE Instrum. Meas. Technol. Conf. 3. 1932-1936 vol.3. https://doi.org/10.1109/IMTC.2001.929537.

Sampaio, D. & Almeida, P. (2016). Pedagogical strategies for the integration of Augmented Reality in ICT teaching and learning processes. Procedia Computer Science, 100, 894-899. https://doi.org/10.1016/j.procs.2016.09.240

Exploring Social Presence in the Metaverse: A User Study on Immersive Interactions

Paolo Barzon, Sylvie Dijkstra-Soudarissanane, and Omar Niamut

TNO, The Hague, The Netherlands

Corresponding authors: paolo.barzon@tno.nl, sylvie.dijkstra@tno.nl, omar.niamut@tno.nl

Introduction

The COVID-19 pandemic has forced many organizations to shift their work lives online. Unfortunately, several aspects of meetings, such as communication and engagement, were negatively impacted. As a result, Immersive Communication Systems (ICSs) have gained increased attention as a potential alternative to the limitations of traditional online meeting platforms.

We research the feasibility of using ICSs for business meetings, driven by the added values they provide, such as increased social presence and effectiveness. We developed an ICS that enables employees from different locations to meet and collaborate in Virtual Reality (VR), providing them with a better meeting experience than traditional meeting platforms.

Starting from traditional online meeting and 2D video communication platforms such as MS Teams[15] and Zoom[16], they are widely used nowadays and denote the most common means of communication between remote users. However, several limitations have been identified that negatively affect the user experience (Bailenson, 2021). Based on our comprehensive analysis of the literature, we have derived a clustering of causes as follows: lower engagement and productivity, worse social interaction and communication, and technical challenges. Our study also highlights previous empirical research demonstrating the effectiveness of ICSs in addressing the shortcomings of traditional meeting platforms. Specifically, the added values of ICSs for business meetings are better communication, enhanced meeting effectiveness, increased product creative quality, the shared real-time interactions, and the positive user perception of the technology and its feeling of immersion.

Our research objective is to assess the viability of implementing ICS as a practical alternative to MS Teams for remote meetings in a business setting. The primary goal of our study is to assess the Quality of Experience (QoE) offered by our ICS, particularly in terms of usability, social communication (comprising spatial and social presence), and overall effectiveness, while drawing comparisons with the well-established MS Teams platform. Our hypothesis is that, despite current potential usability limitations due to technology immaturity, ICSs is more effective than MS Teams in brainstorm meetings, and improves the social communication in meetings involving both two and three participants.

[15] Video Conferencing, Meetings, Calling | Microsoft Teams
[16] One platform to connect | Zoom

State of the art in ICS

The landscape of ICS for business and work applications has expanded significantly over the past years, offering tailored solutions to enhance collaborative experiences. For example, Meta Horizon Workrooms[17] specifically engineered to cater for the dynamics of work meetings. Meta Horizon Workrooms offers a Mixed Reality platform with digitally rendered environments blended with the real world of each user, which can seamlessly access their computer from the Metaverse without having to disconnect. Their shared whiteboard feature allows users to co- create on their products synchronously thanks to a variety of tools, such as images, sticky notes, and digital pens. These features can be also used privately for sketching and private notetaking. When it comes to remote collaboration, an important feature is the capability to share a computer screen, where all participants in the meeting can look at the materials shared by a user. Meta's platform serves as just one example amongst many, such as Glue[18], Connec2[19], and Engage[20], each contributing unique immersive collaboration features. These cutting-edge endeavors are shaping the state of the art in ICS for business and work context, delivering a more effective and engaging remote collaboration to users.

System overview

Our innovative ICS is based on our research on the domain of Social XR, where our human centric designs are delivering meaningful immersive experiences. Unlike state-of-the-art solutions that use virtual avatar representations, our system leverages 3D photorealistic representations to render meeting participants as point clouds in the virtual environment, thereby conveying a better sense of co-presence and realism. Moreover, our system integrates an array of collaborative tools such as screen sharing and virtual presentation screens that enhance productivity and collaborative processes.

Figure 95 shows an example of the system, where two remotely located users are represented as 3D point clouds in a virtual room. The users can see each other in the shared space, which increases the feeling of being together, while the full-body capture enhances the sense of social presence. Figure 96 depicts a schematic representation of three colleagues, connected from separate rooms, experiencing the feeling of sitting together at a table in the VR environment.

Describing in depth the hardware and software components of the TNO ICS is not in the scope of this work. For further research, the paper by Gunkel et al. (2019) explains in depth the web-based VR framework used.

---

[17] Horizon Workrooms virtual office and meetings | Meta for Work
[18] https://www.glue.work/
[19] https://connec2.nl/
[20] Homepage - ENGAGE | The Professional Metaverse & Spatial Computing Platform for Enterprise & Education (engagevr.io)

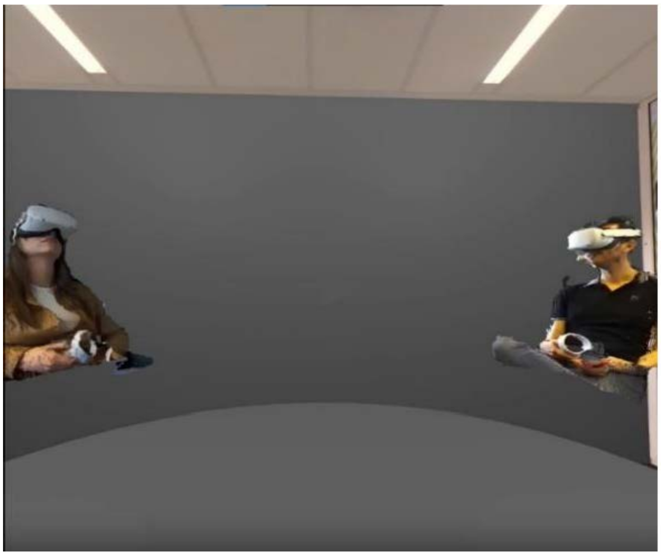


Figure 95. A snapshot of the room in the Metaverse.

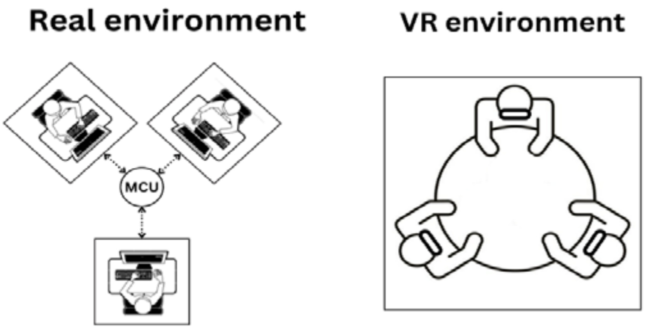


Figure 96. A scheme of the two scenarios

HCI Methodology

Based on insights gathered from the literature review, we developed three Research Questions to assess the feasibility of deploying our ICS. RQ1 explores the changes in the usability of the TNO ICS and MS Teams; RQ2 investigates the different levels of perceived effectiveness across the two communication media, and between brainstorming and broadcasting meeting; lastly, RQ3 examines the variations of social communication in the two communication media, comparing the social and spatial presence in meeting with three and two participants.

To measure usability, we used the SUS questionnaire designed by Brooke (1996). We employed the "Measurement Scales for Perceived Usefulness" designed by Davis (1989) to assess the effectiveness of the systems. Lastly, the H-MSC-Q questionnaire designed by Toet et al. (2022) was used to investigate the spatial and social presence. The interview sessions delved further into the participants' rationale behind their opinions and experience with the systems. The themes of the interviews were usability, effectiveness, and social communication, but did not exclude questions on different topics that could have explained the user's perceptions and behavior.

Main results and achievements

As expected, the results of the SUS questionnaire, on a scale from 0 to 100, showed that there is a significant difference ($p < 0.001$) between the usability of the two communication media. As reported in Figure 3, MS Teams is rated 43 points better than the TNO ICS. According to the SUS score interpretation of

Bangor (2009), the TNO ICS obtained “F”, the lowest of the grades, and an adjective rating of "Ok". For comparison, the grade assigned to MS Teams is "B", which translates to "excellent" in the adjective scale.

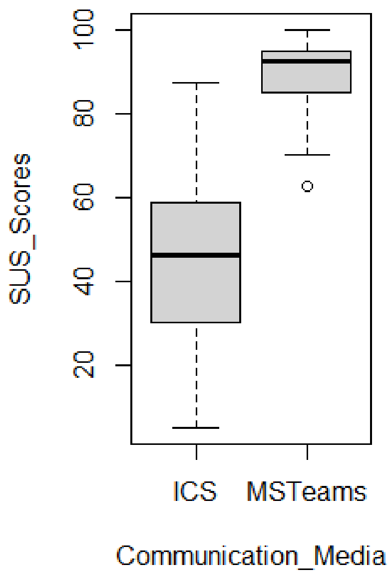


Figure 97. SUS scores of the TNO ICS and MS Teams.

Regarding the effectiveness, when converting the Likert scale of the Perceived Usefulness questionnaire into numbers from 1 to 7, the results show that both brainstorm and broadcast tasks received similar ratings between them, and were both higher for the MS Teams condition. The TNO ICS received better grades for the brainstorm condition compared to the broadcast condition, as expected. The ANOVA results show that there is a significant difference in the effectiveness based on the communication media used ($p < 0.001$). Instead, the meeting type does not show a significant effect on the effectiveness scores ($p > 0.05$). Figure 97 graphically summarizes the data.

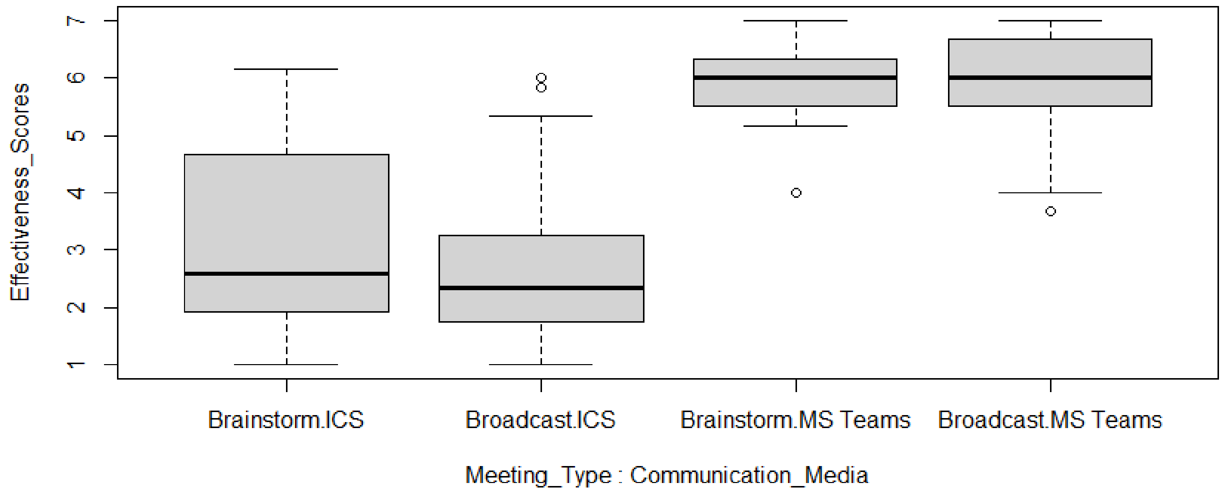


Figure 98. The effectiveness scores of the two variables and their levels.

Lastly, the two sections of the H-MSC-Q investigate spatial presence and social presence. As before, its rankings were translated to a scale from 1 to 7. For both variables, meetings with 3 participants received better ratings than meetings with 2 participants for the TNO ICS, as expected, and for MS Teams too, which instead diverged from expectations. Against the prediction is also the score for the communication media, where MS Teams received better grades than the TNO ICS for both 2 and 3 participants. The ANOVA results show that there is a significant difference in the spatial and social presence scores of the H-MSC-Q based on communication media ($p < 0.001$); however, the number of participants does not have a significant effect on the H-MSC-Q scores ($p > 0.05$). Figure 99 graphically summarizes the data of the spatial and social presence variables.

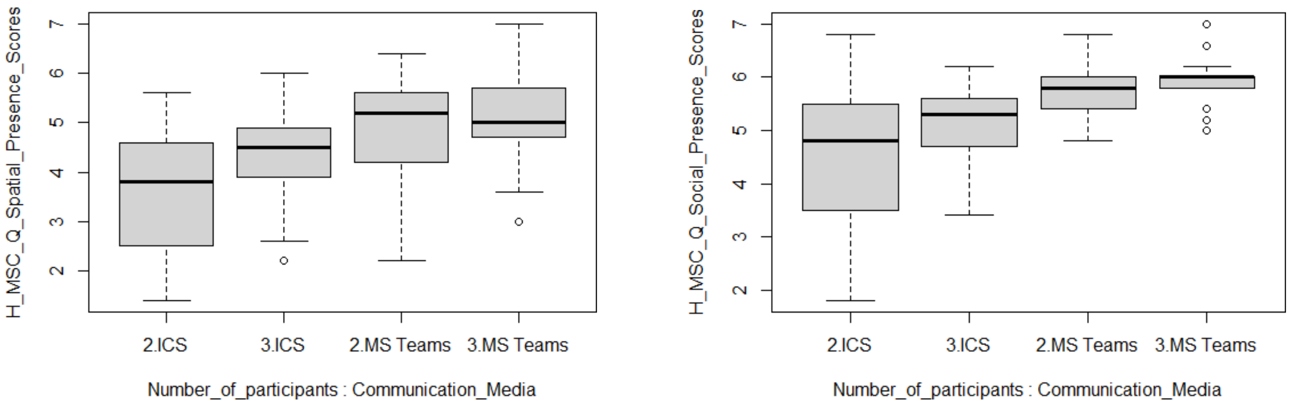


Figure 99. The H-MSC-Q score of spatial and social presence across communication media and participants.

Regarding the qualitative data gathered through the interviews, it emerged that the main issues are related to the HMD covering the partner's eyes, the point clouds being displayed badly or in the wrong position, and the presence of technical issues. The first two disadvantages affect the social communication, while the third is related to usability. As for the suggestions, participants mostly expressed their desire to take notes, interact with the environment, and see their own hands and lower body. The first idea is part of the effectiveness cluster, while the other two are part of the social communication cluster. Lastly, TNO employees appreciated seeing each other in the Metaverse the most, followed by the overall positive experience, and the feeling of being there. All three are part of the social communication group. The summary of the interviews is reported in Figure 100: the codes are divided into the three categories of QoE, namely effectiveness, mediated social communication, and usability; the height of each bar represents the number of interviews in which they appear, out of the 54 analyzed; lastly, a red bar indicates that the code is a complaint, a green bar represents a positive aspects of the ICS, and an orange bar constitutes a suggestion.

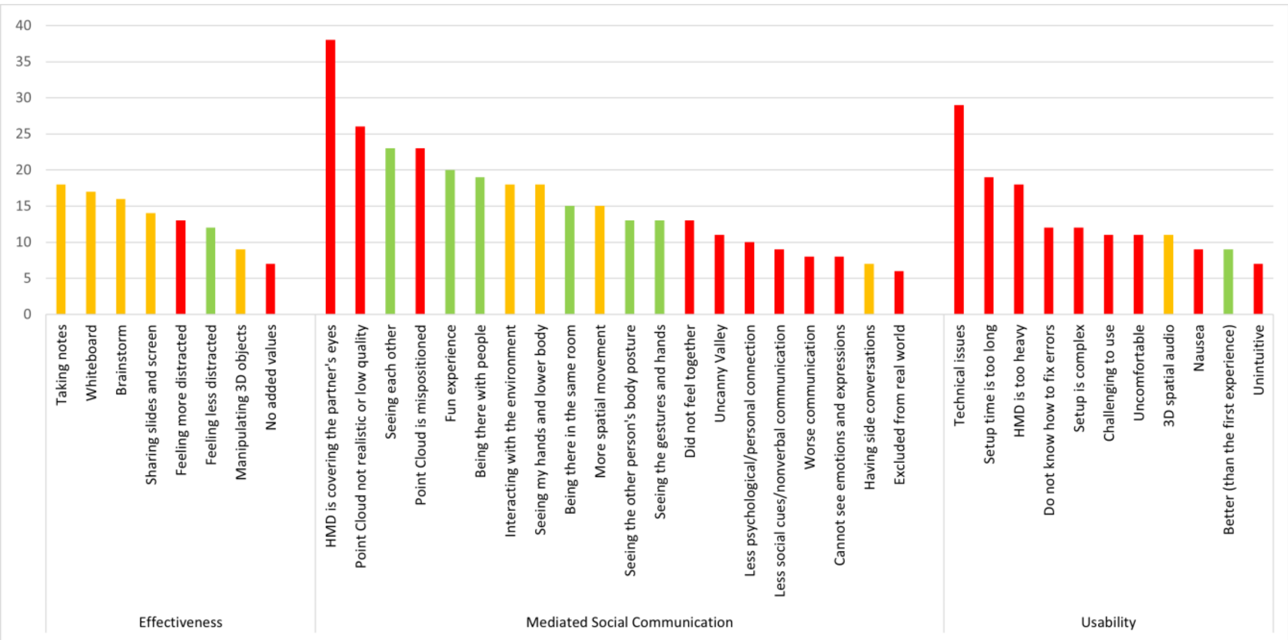


Figure 100. A summary of the codes gathered during the interview analysis.

Conclusion, discussion, future steps

The interviews helped us uncover the main issues with the technology. Firstly, the qualitative data highlighted how usability is the area that hinders the most the overall experience, which is instead enhanced by the social and spatial presence. The sphere of effectiveness, on the other hand, received the most constructive comments for its improvement. As of today, we are already developing a HMD removal technique, and the body capture software is in continuous development; given that those participants who did not mention technical issues or dissatisfaction in the point clouds gave the highest ratings in all three QoE divisions, we expect that fixing the mentioned matters would greatly improve the TNO ICS, making it a valuable meeting platform to be deployed in place of MS Teams. Secondly, the quantitative data showed how the TNO ICS performed no better than its counterpart for any meeting conditions, although providing similar experiences regarding the social and spatial presence. Still, the TNO ICS performs better in brainstorming meetings compared to broadcasting meetings, and in meetings with three participants rather than two.

The results of our study provide valuable insights into the feasibility of using Immersive Communication Systems (ICSs) as an alternative to traditional online meeting platforms, particularly in a business context. Although the usability of the TNO ICS is still insufficient, and technical issues tainted the VR meetings, we are positive that meeting in an immersive space will enhance the meeting experience of employees thanks to its higher levels of spatial and social presence provided.

With our study we hope to pave the way for future enhancements to ICS technology. While the TNO ICS presents a valuable potential in its social and spatial presence provided, further research is required to address the limitations identified in this study and refine the design and implementation of ICSs. The empirical study recognize that although the TNO ICS might not be the best communication media at the moment, the ongoing development of ICS technologies will likely result in more sophisticated platforms that offer enhanced usability, effectiveness, and immersive experiences, for an enhanced meeting experience.

References

Bailenson, J. N. 2021. Nonverbal overload: A theoretical argument for the causes of zoom fatigue. Technology, Mind, and Behavior, 2(1). https://doi.org/10.1037/tmb0000030

Bangor, A., Kortum, P., & Miller, J. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. J. usability studies, 4(3), 114-123.

Brooke, J. 1996. Sus: a "quick and dirty'usability. Usability evaluation in industry, 189(3), 189-194.

Davis, F. D. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, 319-340. https://doi.org/10.2307/249008

Gunkel, S. N. B., Dohmen, M. D. W., Stokking, H., & Niamut, O. 2019. 360-degree photo-realistic VR conferencing. 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). https://doi.org/10.1109/VR.2019.8797971

Toet, A., Mioch, T., Gunkel, S. N., Niamut, O., & van Erp, J. B. 2022. Towards a multiscale QoE assessment of mediated social communication. Quality and User Experience, 7(1), 4. https://doi.org/10.1007/s41233-022-00051-2

Digital Meets Physical – Giving Digital Heritage Objects Presence in a Physical Environment

Monika Keenan

Corresponding author: monika.r.keenan@gmail.com

Keywords: Transparent Display, Digital Heritage Object, Shared AR Experience

Introduction

XR in cultural heritage can facilitate engagement by allowing users to view or interact with heritage sites or objects that may be fragile, or even no longer exist. However, XR experiences do not necessarily facilitate shared experiences, a key motivator for people visiting a heritage site or museum, and the lack of options for embodied interaction with digital objects has been criticised (Jeffrey, 2018; Ciolfi, 2021). Innovative methods for shared displays that unite the physical and the digital may promote better engagement with digital heritage. The Jordanhill Cross at Govan Old Parish Church in Glasgow, UK was chosen as a case study to test the feasibility of a new method of display which not only improves access to digital cultural heritage but promotes a social viewing experience and direct interaction with digital objects.

XR and the Metaphorical Window

To create an innovative XR display, it is important to consider how users might perceive XR experiences. Metaphor theory shows that humans use metaphors in our conceptualisation of abstract ideas. A common concept metaphor employed to help understand some XR systems, AR in particular, is a window (for examples see: Wellner, 2011; Jeffrey, 2018). A window is an opening, through which we can view another space, as is an AR system, therefore the metaphor is: an AR system is a window. This metaphor is understood as follows: a window typically has a glass pane, which allows us to see through, though also acts as a barrier, separating us from whatever lies beyond, like a screen. A window also has a frame, which delineates the view, like the dimensions of a screen. A window is surrounded by a wall, yet another barrier, this time opaque, between the subject and what is viewed, likewise AR systems create a conceptual wall.

Mobile phones, a common device used for AR, can be described as a “wall-window” (Wellner, 2011, p. 81), acting as a metaphorical wall by disconnecting the user from their physical surroundings, while also acting as a window, through which the user is connected to a remote space. It could be argued, however, that the inverse is true for AR itself. AR devices are window- walls, as an AR application opens a window onto an imaginative dimension within the wall of our physical environment. AR technologies mediate our perception of the world by “redirecting” it towards another space or time, showing us a framed view of a dimension that exists somewhere between the material and immaterial (Wellner, 2011, p. 95).

The fact that the world presented in its view is framed, limited, and somehow barred from us encourages an imaginative journey (Kenderdine, 2007; Jeffrey, 2021). In our mind's eye, we might pass through the window glass into the time or space that lies beyond (Jeffrey, 2018). This effect is facilitated by the transparency of the technological device being used. Here, transparency refers to whether the mediation device draws our attention while in use. The moment we must become active in deploying the technology or see it working, it becomes harder to move through the glass (Liberati, 2013). Importantly, though, there is an "ontological cut" marked by the frame of the screen, which distinguishes the metaphorical wall from the view through the metaphorical window (Friedberg, 2006, p. 5). This refers to the distinction between the materiality of the screen and the immateriality of the view within it. The screen itself represents a barrier made of glass that cannot be crossed (Jeffrey, 2018), a concept which highlights the issues surrounding embodiment encountered by many XR systems.

Transparent Screens for XR Displays

One of the affordances of transparent displays is that they combine the environment and augmented images or information in a common perceptual space (Queisner 2017; Friedberg, 2006). The fact that digital heritage objects seem to lack a materiality can disrupt the way people engage with them, even potentially fostering a sense of disconnection with the object (Jeffrey, 2015). A transparent screen, then, which allows the digital and the physical to inhabit the same space for us perceptually, might reestablish a link between a physical object or space and its digital replication or augmentation. Transparent screens offer a more direct connection between the viewer and what is being viewed by improving our sense of proximity with the object and facilitating the creation of a space in which the physical and the digital are concomitant, and therefore promote engagement with digital heritage. Furthermore, the appeal of a transparent screen for creating such a hybrid space is evinced by the enduring use of the Pepper's Ghost technique. This technique, which uses mirrors, tilted glass and lighting to create the illusion of an apparition, dates back to the 16th century and is still used to great effect at theatres and concerts today (Brooker, 2007).

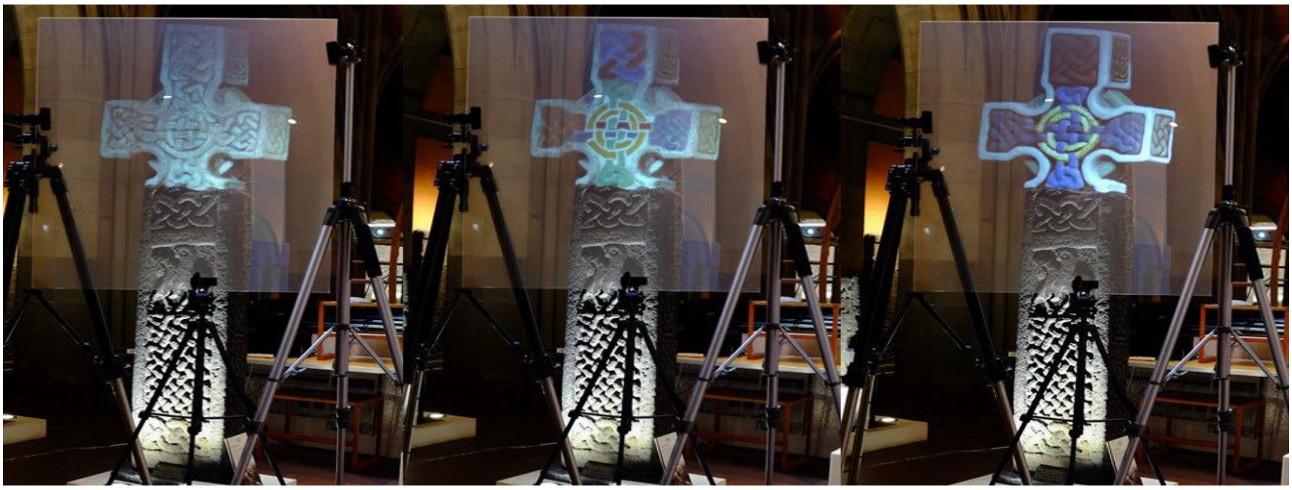


Figure 101. *An animation of three visualisations of the crosshead displayed as an installation.*

Creating and Testing a Transparent Display

To determine the benefits of using transparent displays, a display device was designed and created. As discussed, the frame plays an important role in how AR applications are perceived. If we remove the visual frame, therefore, we may be able to remove a conceptual barrier to the digital heritage object by integrating it into the physical world seamlessly. The original design of this device removes the frame and uses wooden feet to balance and support a piece of acrylic sheet. With no frame being in the visual field at eye-level, the illusion of the material and digital inhabiting a joint perceptual space is not disrupted by the conceptual barrier of the frame. However, due to the limitations of the project, it was not possible to create a full-size version of the original design. Instead, a prototype was created. Rear projection foil was applied to a piece of acrylic sheet which was held in place with tripods and clamps to assess the performance of the device on site.

Figure 102. Left: original design for AR device. Right: Temporary installation set up on site at Govan Old.

The Jordanhill Cross, one of the Govan Stones, was chosen as a case study and an installation of visualisations was created. The Jordanhill Cross is an early medieval free-standing cross, though the crosshead is missing and only the carved shaft remains. However, a concrete replica stands outside the church. Using photogrammetry, three digital visualisations of the complete Jordanhill Cross were created and projected onto the display device to create an installation of digital visualisations, through which the existing stone base is still visible in situ. Multiple visualisations, in this case, digital colourisations, were displayed to better communicate the uncertainties inherent in heritage interpretation. Physical and digital elements were combined in a shared space to recontextualise hypothetical historical possibilities.

Figure 103. Three digital visualisations of the Jordanhill Cross-head produced using photogrammetry.

Outcomes

The installation on site showed that this method of display was effective, despite the use of a prototype rather than a finished version of the original design. Importantly, unlike projection mapping, this method allows for the projection of objects, or parts of objects, that do not materially exist. It was also widely visible in the space, allowing and even encouraging a shared experience. While the device was static, and the AR experience could only be viewed from a fixed position, it nevertheless effectively combined a physical object with its digital augmentation in a joint space and gave the digital visualisations a striking sense of presence in the physical space. Viewers found the augmentation on a transparent screen helped bring the original stone to life and created a dynamic experience. However, there is certainly scope for development and future improvements in this concept. The projected augmentation was only a 2D image. While testing the device on site, though, this did not appear to minimize the impact of the installation with audiences, and the benefits of this method, such as the low cost of the design, reduced barriers for use and improved opportunities for a social experience seem to outweigh this drawback. The project demonstrates that this is a feasible method for creating a shared AR experience, and that the use of transparent screens can help to improve engagement with digital heritage objects by reestablishing a link with the physical environment.

References

Brooker, J. 2007. The polytechnic ghost: Pepper's ghost, metempsychosis and the magic lantern at the Royal Polytechnic Institution. Early Popular Visual Culture, 5(2), 189-206. https://doi.org/10.1080/17460650701433517

Ciolfi, L. 2021. Hybrid Interactions in Museums: Why Materiality Still Matters. In Champion, E. M. (ed.) Virtual Heritage: A Guide. London: Ubiquity Press, pp. 67-79.

Friedberg, A. 2006. The Virtual Window: From Alberti to Microsoft. Cambridge, MA: The MIT Press.

Jeffrey, S. 2015. Challenging Heritage Visualisation: Beauty, Aura and Democratisation. Open Archaeology, 1(1), 144-152. https://doi.org/10.1515/opar-2015-0008

Jeffrey, S. 2018. Digital heritage objects, authorship, ownership and engagement. In Di Giuseppantonio Di Franco, P., Galeazzi, F. and Vassallo, V. (eds.) Authenticity and cultural heritage in the age of 3D digital reproductions. McDonald Institute Conversations. Cambridge: McDonald Institute for Archaeological Research, University of Cambridge, pp. 49-56.

Jeffrey, S. 2021. The work of the miniature in the age of digital reproduction. In Dawson, I., Meirion Jones, A., Minkin, L. and Reilly, P. (eds) Diffracting Digital Images: Archaeology, Art Practice and Cultural Heritage. Routledge. https://doi.org/10.4324/9781003042129-6

Kenderdine, S. 2007. Speaking in Rama: Panoramic Vision in Cultural Heritage Visualization. In Kenderdine, S. and Cameron, F. (eds.) Theorizing Digital Cultural Heritage: A Critical Discourse. Cambridge, MA: The MIT Press, pp. 301-331.

Liberati, N. 2013. Improving the embodiment relations by means of phenomenological analysis on the "reality" of Ars. In IEEE International Symposium on Mixed and Augmented Reality 2012 – Arts, Media, and Humanities (ISMAR-AMH), pp. 13-17. https://doi.org/10.1109/ISMAR-AMH.2012.6483983

Queisner, M. 2017. Disrupting Screen-Based Interaction - Design Principles of Mixed Reality Displays. In Busch, C., Kassung C. and Sieck, J. (eds.) Kultur und Informatik: Mixed Reality. Glückstadt, Germany: Verlag Werner Hülsbusch, pp. 133-144.

Wellner, G. 2011. Wall-Window-Screen: How the Cell Phone Mediates a Worldview for Us. Humanities and Technology Review, 30, pp.87-103.

Augmented Reality for Developing Interior Design Skills in Secondary Education

Zinta Zalite-Supe, Lana Franceska Dreimane, Anna Ansone, Linda Daniela, and Astra Rudolfa

Faculty of Pedagogy, Psychology and Art, University of Latvia, Rīga, Latvia

Corresponding author: zinta.zalite-supe@lu.lv

Keywords: Augmented Reality, Competences, Effectiveness, Interior Design, Secondary Education

Introduction

Practitioners and researchers in the field of interior design agree that modern interior design education is not conceivable without the use of immersive technology, as this enables educators to utilise its potential to address the ever-evolving demands of interior design as a field. The main aim of this study is to analyse the framework of "Design and Technologies" subject and map out the pathways for the integration of AR with the defined learning objectives in order to develop interior design skills. Since 2020, Latvian primary and secondary education have shifted towards a competence-based approach and curriculum, including the subject "Design and Technologies" that covers interior design. This paper presents a summary of the analysis of ten AR learning experiences that could potentially be applied in the learning process of "Design and Technologies" subject, particularly focusing on the interior design module. AR learning experiences were analysed using a VR learning experience evaluation tool developed by Dreimane (2020). In order to compare the effectiveness of the learning process and the experience of students, the authors validated the most highly evaluated learning experience with twenty-two secondary school students of 10th grade (aged 16 years) from one school in Latvia. In addition, survey data were collected in order to gain a deeper understanding of students' experience. Furthermore, this study offers practical guidelines (for educators) for developing interior design skills at a secondary school level, based on lessons from the validation of the AR learning experience. In addition, a summary of platforms where educators could develop content in AR was created.

Methodology

This research employs a design-based study approach (Brown & Collins, 2018) and various qualitative research methods, involving six steps:

1. Conducting a literature review on AR benefits and potential risks in interior design education.
2. Creating a framework for integrating AR to develop skills aligned with the interior design module included in "Design and Technologies" subject.
3. Validating the framework through expert interviews.

4. Analyzing AR learning experiences and content development platforms that have the potential to be employed in the learning process of "Design and Technologies".
5. Testing and validating the highly evaluated AR learning experience and developed framework with a group of 10th-grade students.
6. Categorizing platforms suitable for educators to create AR content for teaching interior design.

The framework for integrating AR into interior design skills, aligned with the "Design and Technologies" subject, was validated through interviews with two educational experts and three educators from various Latvian secondary schools in March and April 2022. The study also used the "Nodes" method (Robins, 2015) to analyze AR learning experiences and content development platforms. The study tested and collected data on twenty-four applications, including ten AR learning experiences, eleven content development platforms, and three hybrid applications, to assist educators in teaching interior design. An adapted VR evaluation tool by Dreimane (2020) was used to assess ten pre-selected AR learning experiences. As the content development platforms have a different purpose from learning experiences, they were evaluated according to the criteria defined by authors: (1) technological criteria; (2) pedagogical criteria; (3) user experience criteria. The AR learning experiences which were analysed are following (Figure 104.):

1. "MytyAR";
2. "Live Home 3D";
3. "Houzz";
4. "IKEA Place";
5. "Decormatters";
6. "Planner 5D";
7. "Measure";
8. "Rooomy";
9. "Magic Plan";
10. "Dulux".

The AR content- development platforms which were analysed are following:

1. "Three.js";
2. "ZapWorks";
3. "UniteAR";
4. "AdobeAero";
5. "ViewAR";
6. "Turbosquid";
7. "Sketch-Up";
8. "Vuforia";
9. "Unity 3D";
10. "Blender";
11. "Unreal Engine".

The remaining three applications: 1) "Class VR"; 2) "Live home 3D"; 3) "3DBear", contained both learning experience and development possibilities.

Myty AR Decormatters Planner 5D

Figure 104. The results of applications testing: "Myty AR", "Decormatters", "Planner 5D"

The highest-rated learning experience was validated at a secondary school with a group of 10th-grade students (16 years old, twelve female, and ten male participants). The study compared the effectiveness of the learning process and student experience, validating the framework through two comparative lessons using different approaches to foster the same skills. Based on lessons from the validation of the AR learning experience, this study provides practical guidelines (for educators) to foster interior design skills at the secondary school level. In the first class, the assignment was completed using traditional materials. Prior to this assignment students received a theoretical introduction to lighting and various materials applicable in interior design, accompanied by a diverse range of examples. Students were tasked with creating an interior design collage, followed by argumentation and description of the chosen conceptual mood, utilization of the materials, and the significance of lighting. The collage was constructed using images from interior, architecture, and design magazines, along with colored pencils and paper. This task was undertaken in pairs. Subsequently, one week later, during the second class, the same assignment was executed by utilizing AR, by applying the highest regarded learning experience titled "Live Home 3D", which were evaluated by authors based on a VR evaluation tool (Dreimane, 2020). A student survey was conducted, categorizing their responses into three groups: (1) benefits of AR technology use in learning; (2) assessment of the learning process from student experience; and (3) conclusions on AR technology's lesson effectiveness.

Results

The literature analysis revealed five main benefits for AR implementation in curriculum: 1) students can experiment, make mistakes, and learn in real-time (Liono et al., 2021). 2) AR allows convenient use of materials, lights, colors, and spaces compared to slower and time-consuming traditional methods like paper and pencil, which improves contextual understanding (Buchner et al., 2021). 3) Learning interior design through AR experiences can assist students in preparing for the evolving work environment, while also acquiring knowledge that will establish a robust foundation for subsequent educational stages (Chang et al., 2019). 4) When information is visualized using AR, the knowledge gained tends to stay fresh in memory for an extended duration. In today's rapidly changing information landscape, it's crucial for effective learning to encompass a comprehensive grasp of educational content and the ability to retain this understanding over a prolonged period, as emphasized by Buchner et al. in 2021. 5) AR technology is emerging as a crucial tool for prototyping due to its ability to link the physical environment with the virtual environment through interactive three-dimensional models. This integration allows for a seamless interaction between the physical surroundings and the virtual elements. Consequently, AR technology facilitates the creation of prototypes that provide a more realistic and immersive experience for students. This process, in turn, fosters creativity (Chang et al., 2019).

"Design and Technologies", is a completely new subject area that requires educators to familiarise themselves with the curriculum framework and effectively employ a range of new methods and tools to facilitate a high-quality learning experience. The aim of this study was to develop a framework of mapped-out pathways for the integration of AR in the development of skills defined by "Design and Technologies" and to validate the results with experts and educators of "Design and Technologies" subject. Interior design is taught through finding answers to three pivotal questions: 1. How? Plan and study the process of creating design solutions; 2. With what? Choosing Materials and Technologies; 3. Why? Creating a design solution (justification for the concept) (The National Centre for Education, 2020). Each pivotal question was addressed through a mapped pathway consisting of the following sections: Skills (aligned with the curriculum), Rationale for AR implementation, and Applying AR to develop specific skills. The framework is accessible through this link: https://b2ty.short.gy/framework.

Framework was validated through expert interviews. Experts and educators confirmed the framework's practicality for AR technology in interior design, but suggested adding AR learning experiences and content-building platforms for classroom use. The experts pointed out that the mapped-out framework clearly shows skills where AR technology can be applied. In order to provide educators with practical tools for teaching interior design, this research conducted testing and data collection of twenty-four applications. One of the key lessons that emerged is that there are relatively few platforms designed specifically to be used in the learning process of developing interior design skills. Out of a total of ten selected AR learning experiences, only two directly focused on learning ("Live Home 3D" and "Planner 5D"). The most suitable AR for classroom integration is "Live Home 3D", offering students enhanced versatility in developing interior design skills through a heightened sense of presence in an updated reality environment. This experience overlays digital information onto existing interiors, facilitating a natural and lifelike learning extension. It also provides video-based instructions for interior design and tool usage. Evaluation confirms "Live Home 3D" promotes higher cognitive skills like creation, evaluation, and development within Bloom's taxonomy. Additionally, it enables simultaneous design solutions for entire spaces, supporting experimentation and mood creation.

Further, the learning experience "Live Home 3D" was validated with 10th Grade students (Figure 105).

Figure 105. Some of the students' projects in the application "Live Home 3D"

The study highlights the benefits of using AR technology in the learning process. 91% of students reported gaining versatile knowledge in interior design, including materials, 3D room setups, lighting, and more. Students identified three key advantages of AR: 1) practical and diverse experimentation; 2) engaging learning through visualization, and 3) increased creativity through collaborative work. Students supported AR integration, with 95% expressing positive feedback. They found AR beneficial for understanding room dimensions and connecting interior design to real-life scenarios, emphasizing its effectiveness.

Conclusions

The aim of this study was to develop a framework of mapped-out pathways for the integration of AR in the development of skills defined by "Design and Technologies" and validate the results with experts and educators of the subject. Based on the research results, using AR in the learning process has been more effective than traditional materials. The data essentially corroborate the claims made in the literature analysis of this current study, providing a justification for the integration of AR technology into the curriculum of Design and Technologies. This is also reflected in the replies provided by the secondary school students themselves, experts and an examination of AR learning experiences. In order to effectively integrate AR technology for the development of interior design skills at the secondary schools level: 1) it is necessary to define a plan to be achieved by AR technology; 2) the educator must have knowledge of the role of AR technology in the learning process (including related competencies), as well as knowledge of the aspects of organising the learning process in a technology-enhanced environment; 3) and an assessment of expected results, benefits and risks must be carried out; the assessment must provide a clear vision of: 4) whether AR technology can be integrated into the learning process; and 5) whether AR learning experience enhances the planned learning process. Future research objectives include long-term validation of AR's effectiveness, creating country-specific AR learning experiences based on research findings, and exploring AR integration into other modules of the "Design and Technologies" subject's curriculum beyond interior design module.

References

Brown, L.A., & Collins A. 2018. Design-Based Research. Available: https://methods.sagepub.com/reference/the-sage-encyclopedia-of-educational-research- measurement-and evaluation/i7087.xml?fromsearch=true

Buchner, J., Buntins, K., & Kerres, M. 2021. A systematic map of research characteristics in studies on augmented reality and cognitive load. Comput. Educ. Open 2, 100036. https://doi.org/10.1016/j.caeo.2021.100036

Chang, Y. S., Hu, K. J., Chiang, C. W., & Lugmayr, A. 2019. Applying Mobile Augmented Reality (AR) to teach Interior Design students in layout plans: Evaluation of learning effectiveness based on the ARCS Model of learning motivation theory. Sensors, 20(1), 105. https://doi.org/10.3390/s20010105

Dreimane, L.F. 2020. Virtual reality learning experience evaluation tool for instructional designers and educators'. In L. Daniela (ed.), New Perspectives on Virtual and Augmented Reality: Finding New Ways to Teach in a Transformed Learning Environment, Taylor&Francis, pp. 3-21.

Liono, R.A., Amanda, N., Pratiwi, A., Gunawan, A.A. 2021. A systematic literature review: learning with visual by the help of augmented reality helps students learn better. Procedia Comput Sci 179, pp. 144-152. https://doi.org/10.1016/j.procs.2020.12.019

Robins, G. 2015. The difference with social network research. In: Doing Social Network Research, SAGE Publications Ltd, pp. 1-16.

The National Centre for Education. 2020. Competence Approach to The Teaching Content of Learning, Available: https://mape.gov.lv/api/files/001087FE-759D-4AB4-A05F-D9F714B6B4FE/download, ISBN 978-9934-540-96-7

Mixed Reality Collaboration in Industrial Scenarios

Marton Szabo-Kass, Michael Kernbichler, and Daniel Fabry

Institute for Design and Communication, University of Applied Sciences, FH JOANNEUM Graz, Austria

Corresponding authors: marton.szabo-kass12@fh-joanneum.at, michael.kernbichler@fh-joanneum.at

Keywords: Collaboration, Mixed Reality, Virtual Interaction, Product Development, Industrial Production, Collaborative Training, User Interfaces, Serious Game

Introduction

Online collaboration is a basic component of everyday work life in many different branches. With the evolution of mixed reality hardware, the advantages of three-dimensional immersion can be utilized in collaborative working scenarios (Wang et al. 2020, 1718). Several applications have been developed for virtual co-work in recent years, but the field is still missing widely established platforms and specific design standards or guidelines.

Within the "ICON – Immersive Co-Creation Hub" research project at the FH JOANNEUM University of Applied Sciences in Graz, Austria, several specific showcases are being developed for demonstrational and research purposes. They are designed to facilitate synchronous and asynchronous mixed-reality meetings of local and remote participants involving real-life equipment and virtual models via mixed-reality headsets. The aims are to explore, analyze and test the interaction experience in shared spaces in industrial scenarios. The findings and lessons learned should be summarized in a toolbox and design guideline collection.

Collaboration in Mixed Reality Environments

Online collaboration dynamics are known from video conferencing platforms, but more research and exploration still need to be done on 3D mixed reality (MR) collaboration. With developing the showcase scenarios, the necessary elements and their effects are investigated and probed with usability tests.

For the functionality and inclusivity of collaborative sessions in shared spaces, the participants joining from remote locations must be represented visually in 3D. Instead of live camera streams known from 2D platforms, 3D avatars are commonly used at different fidelity levels. For the ICON showcases, realistically formed cartoon "Upper Body" avatars (Schäfer et al. 2022, 11) with head, chest and hands were used. The conducted user tests verify these avatars' sufficient functionality, as Yoon et al. also described (Yoon et al. 2019, 554), except for missing facial expressions, that could reflect moods and emotions, as Schäfer et al. stated as well (Schäfer et al. 2022, 16). Simultaneously, it must be considered that participants joining from the same room do not need to see each other's avatars. These would only overlay on their actual body and hinder a seamless interaction with virtual tools and objects. Figure 90 illustrates a possible setup of several spaces with different users and hardware setups, connected within the application.

Audio connection is an essential requirement, ensuring the natural co-presence experience. After getting familiar with the setup, talking face-to-face with an avatar was found to work acceptably. For a more realistic immersion, spatial sound is necessary to recognize the participants' location. This feature is mainly dependent on hardware capabilities though.

To facilitate collaboration, basic information about the participants, such as clear identification and applicable status, must be provided. As exemplary solutions, the showcases were designed with information bars above the avatars, containing real photos, names, roles and audio status indicators.

To make collaboration and immersion in a shared space possible, spatial anchors were utilized to ensure spatial alignment across multiple devices. Such alignment was crucial in achieving congruence between the real and virtual environments. Tests showed that even the slightest spatial deviation can irritate users.

Mixed Reality Interaction

After establishing the basic requirements for collaboration, interaction occurs around the topic and objects of interest. In one of the showcases of the ICON project, an illustrative industrial scenario was developed. It includes an assembly workbench with a collaborative robotic arm (cobot) and parts of an electric scooter wheel motor to be assembled. These objects exist both in a real setup and as virtual 3D models, with which a real industrial assembly process can be developed and trained in an MR setting. The scenario allows the exploration of interaction with virtual 3D objects in space and at the real workbench with the cobot. Additionally, holographic user interfaces show information, instructions or controls of the scene or the robotic arm. Figure 91 shows this scenario in a hybrid session with the real workbench plus a robotic arm, virtual objects, a local participant and a remote participant represented by an avatar. Figure 92 shows a solely local collaborative session without remote participants and with a completely virtual workbench setup.

The objects and interfaces can be manipulated in mixed reality with the help of controllers or only by hand gestures, depending on the hardware used. The Microsoft HoloLens 2 supports hand tracking and gestural interaction in this project. Therefore, the virtual elements can be grabbed and operated intuitively by hand. The conducted user tests showed that visual and audio feedback of all actions contributes to a fluent and transparent experience. For example, audio signals of UI operations or task completions help the user's understanding of tasks and interactions. Also, visual status indications of grabbed objects can prevent misunderstandings between users. Unintended parallel actions could still occur though, that need other solutions.

In the research, user-friendly cobot controls were incorporated to increase intuitiveness. One notable feature is the ability to teach the collaborative robot new positions in the physical world by merely relocating the robotic arm's gripper through manual manipulation. The virtual twin of the cobot then follows these real-life instructions. Additionally, this functionality is bidirectional, enabling the manipulation of the real cobot from mixed reality.

When mixing the real world with virtual components, some restrictions and guided controls could be necessary to maintain structure and prevent unintended actions. One solution is the use of virtual tokens, which ensure that interactive objects can only be edited by one person at a time. Further, participant roles were introduced to control general permissions for interaction with the virtual elements. For the ICON scenarios, the roles of host, editor and viewer were implemented. The host can interact with everything and edit other user's roles and settings, while editors can only interact with the shared objects. Viewers on the other hand are not capable of interacting with these objects at all, they only have an observational role. The effectiveness of this concept was also proven by Pouliquen-Lardy et al., 2016.

Some objects, such as UI elements and interfaces, are not visible to all participants nor shared in the virtual session – everyone sees and can interact with their own copy of the UIs. This led to one of the challenges encountered during the testing phase: users failed to recognize when other participants were occupied with their own user interface. A unified "ghost tablet" was introduced to address this issue, always appearing in front of the users, signaling their engagement.

A novel feature was created to incorporate annotations and photo/video recordings, streamlining the documentation of collaborative sessions. The annotation function enables the inclusion of alerts and indicators within the scenes. At the same time, the photo and video features allow for the comprehensive documentation of each step of the simulated production process. These recordings can then be utilized for post-processing and developing the training scenario.

Current Status and Future Work

At present, two exemplary showcase prototypes are being developed with Unity[21]. The first usability tests have already been conducted and will be performed again at advanced stages. After completion, the gained insights will be structured and documented on a wiki-like online platform, allowing future additions. This should include all relevant components and their application details as well as design guidelines, derived from all lessons learned throughout the development and testing.

Furthermore, the application will be ported to other mixed-reality headsets besides the Microsoft Hololens 2. Thanks to the OpenXR (OpenXR 2023) standard, there are several possible platforms (Meta Quest Pro, Varjo XR-3) to test the showcases further. Afterward, the application's simplified, spectator-only version will be exported for tablet use.

[21] Unity Engine. Unity Technologies 2023. Available at: https://unity.com/

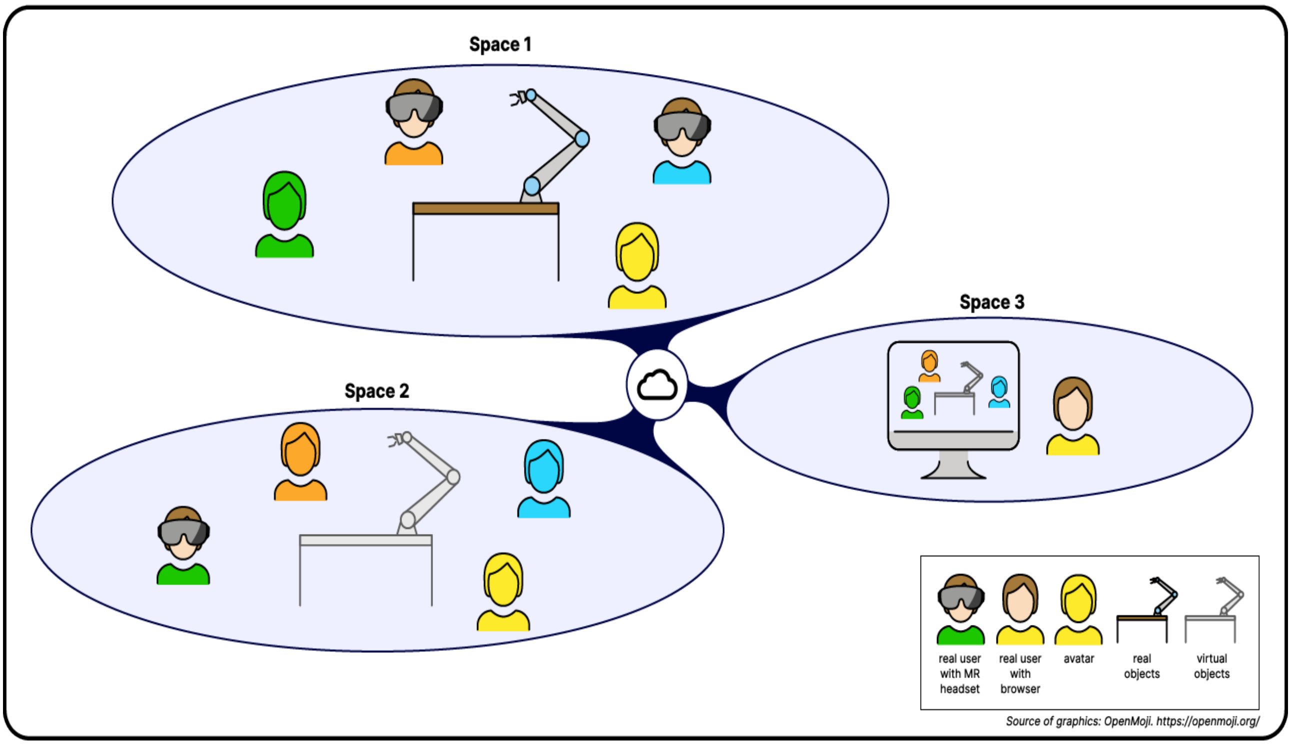


Figure 106. Illustration of connected spaces with different users and hardware setups

Acknowledgements

The project "ICON – Immersive Co-Creation Hub" is co-funded by the Austrian Research Promotion Agency (FFG), within the 8th call for U.A.S. – Research for the Economy (project code: 884103).

Figure 107. Industrial showcase with real workbench plus robotic arm and virtual objects in a hybrid session

Figure 108. Local session with only virtual elements

References

OpenXR. The Khronos Group Inc. 2023. Available at: https://www.khronos.org/api/index_2017/openxr Pouliquen-Lardy, L., Milleville-Pennel, I., Guillaume F. & Mars F. 2016. "Remote Collaboration in Virtual

Reality: Asymmetrical Effects of Task Distribution on Spatial Processing and Mental Workload." Virtual Reality 20 (November 1, 2016). https://doi.org/10.1007/s10055-016-0294-8.

Schäfer, A., Reis, G. & Stricker, D. 2022. A Survey on Synchronous Augmented, Virtual, and Mixed Reality Remote Collaboration Systems." ACM Comput. Surv. 55, 6, Article 116, 27 pages. https://doi.org/10.1145/3533376

Wang, P., Zhang, S., Billinghurst, M., Bai, X., Weiping H., Wang, S., Sun M. & Zhang, X. 2020. A Comprehensive Survey of AR/MR-Based Co-Design in Manufacturing. Engineering with Computers 36, 1715–1738. https://doi.org/10.1007/s00366-019-00792-3.

Yoon, B., Kim, H., Lee, G. A., Billinghurst, M., & Woo W. 2019. The effect of avatar appearance on social presence in an augmented reality remote collaboration. In Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (VR'19). IEEE, 547–556. https://doi.org/10.1109/VR.2019.8797719

Effects of Extended REALITY Visuals on Game Playability through User Satisfaction Level

Meriam Elleuch[1], Yassine Aydi[1], and Mohamed Kchaou[3]

[1]*Higher Institute of Arts and Craft in Sfax, CES Laboratory (ENIS), University of Sfax, Tunisia.*
[2]*Department of Industrial Engineering, College of Engineering, University of Bisha, KSA, Saudi-Arabia*

Corresponding author: yassine.aydi@enis.tn

Keywords: XR Games, Playability, User Satisfaction, Visual Elements

Abstract

Extended reality (XR) applications, highly useful for the design of immersive experiments in several domains, are on the cutting edge of technological innovation. It is in this context that the present work can be set, with the aim of investigating the Game Design area, jointly explored in collaboration with the Mixed Reality and Visualization (MIREVI) domain specialized team. This study investigates the importance of XR graphics in game creation by extracting how the visuals affect the game's playability by analyzing the user's satisfaction level while playing. We have conducted a user observation during gameplay, and analysis of user preferences over several gaming sessions to determine the level of happiness in a video game. It appears that the unreal sensation and satisfaction level are most influenced by the graphic quality, followed by the idea and the game's style. Through our study, we have concluded that Playability (P) is the amount of sensation level (SL) times the visual quality (VQ). After evaluating the effect of visual components on the users' satisfaction level, some expert methods have been applied to confirm the affirmations of our creation. The results show that XR visuals have a way of mixing authentic life visuals with virtual-created ones.

Visual elements level-related playability

The gaming experience begins when the users are motivated to play, and the centre of users' satisfaction is when they are stimulated. Moreover, motivation is a decisive factor that encourages others to achieve their objectives. It is a stimulus to accomplish something that is motivation. In a precise definition, to be moved to action is to be motivated (Cheah et al., 2022). Motivation is the centre of related game aspects connected to the user's satisfaction level. The user's satisfaction is the main objective of any game design, and to reach a high level of satisfaction, several profound research and work must be followed. In this paper, the process seen in Figure 109 has been followed to create and make a complete study of VR game design. The user experience usually focuses on deeply understanding the user's needs, values, abilities, and limitations.

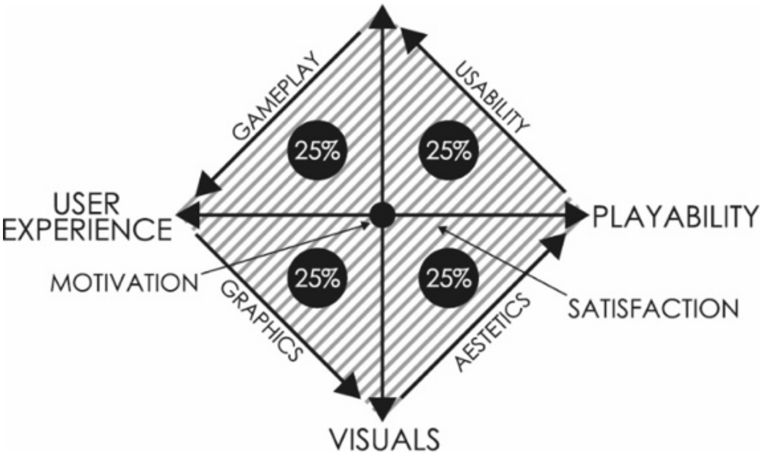


Figure 109. Game-Related User Stimulation

The user interacts with graphics in games; they are the bridge between the gameplay and the player; graphics are one of the essential elements of a video game, by reaching the user's satisfaction through his engagement during gameplay and creating a better visual engine by implementing better texture, adding real effects, faster user movement, and precise details of virtual objects. These contributions increase the realism level in the game's virtual environment (Lrez, B. M., 2014). PLAYABILITY is a set of properties describing the user/player experience using a specific game system/ video game whose main objective is to provide enjoyment, entertainment, and learning strategies. It is related to the game through USABILITY, which presents the game mechanics and controls (Sánchez, José Luis González, et al., 2009). Due to the numerous factors that may be considered in video games, including mental, emotional, physical, entertaining, and social aspects, satisfaction is essential. Employing surveys, player observation during gameplay, and analysis of user preferences over several gaming sessions, the level of happiness in a video game is calculated. In Figure 110, the percentage of users' satisfaction attributable to gaming is divided into four categories according to how it affects them.

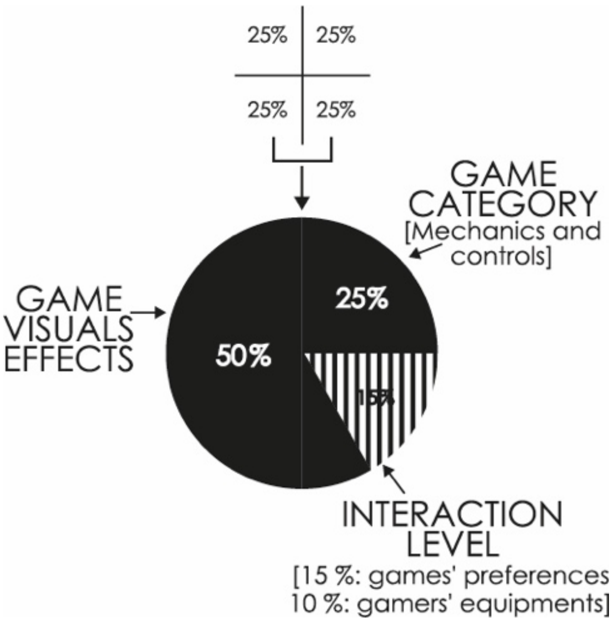


Figure 110. Gamer's satisfaction level

In the context of a video game's usability, satisfaction refers to how pleased users (players) are with that environment. Several qualities, including amusement, beauty, drive, emotion, and sociability, are considered in this measure. The obtained satisfaction rating is an essential quality when discussing video games. Three specific qualities make up this trait. First, the game concept, the game's story, and What defines the game idea are shown in Figure 111.

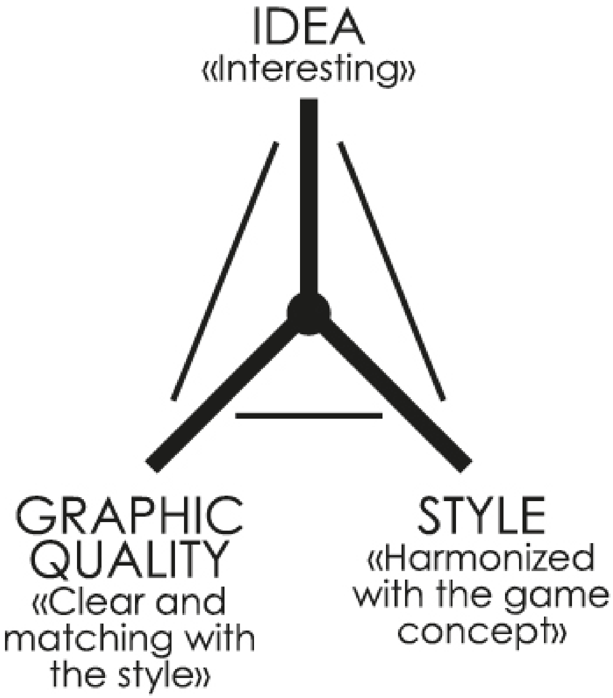


Figure 111. Visual Elements

The stages, objectives, and even the character's role explain the game's idea. Its aesthetic characteristics are more alluring the more intriguing they are. Second, graphic quality has been recognized as a result of advancements in digital technology. The development of visual qualities has progressed from pixels to high-quality graphics; the game's graphical quality influences how the user perceives the game environment; as long as it is simple to interact with, its visual features are appealing. The third characteristic of a game's visual components is that the graphics must complement the game's aesthetic. A game's presentation to the user is referred to as its style (Egenfeldt et al., 2013). The three types of gaming styles are realistic, stylized, and abstract. These three factors determine the visual excellence of the game. The game's visual elements significantly influence the user's overall experience since they evoke strong emotions in the player.

Unreal sensation in XR games

Experimental research was done to determine how the visual design affects the player's experience and the game's playability. The study was located at a lab associated with the University of Applied Science's Department of Media by a MIREVI (Mixed Reality and Visualization) team member. Unreal sensations in XR games, such as teleporting and size changing, are not usually felt in these games. Perceiving the virtual character changes or moves from one place to another on a screen is ineffective for the user. We have added these sensations in VR games to verify how the graphic quality affects the user and higher satisfaction in the game and its playability. We have picked the abstract style of the game and VR graphic quality. Only the idea of these games and the given sensations are different; we came up with three game concepts, each offering a unique experience to the user. Our goal in creating three VR games was to investigate how the user feels while engaging in various gameplay scenarios and how the three vital visual components, mainly the "idea" of the game, affect how satisfied the user is with their interactions in the virtual world. We thought of implying unusual sensations to the human brain to extract how much these visuals impact the user's emotions.

To better study the graphical part of the XR game's effect on the user's mind, we chose an idea that provides the user with a different unreal sensation, such as teleportation. This idea is creating a maze game and using simple interaction between the user and the virtual object, especially for those trying VR gaming for the first time. These ideas have brought many other fascinating insights to follow in our study of games' playability; the idea was a collection of different thoughts from our previous ideas. It is divided into two graphical styles, seen in Figure 112 and Figure 113; the first is an abstract style, seen on the left, and the second is stylized, seen on the figure’s right side. By choosing the idea of creating a simple VR maze game, we started collecting the proper steps to design such a game, starting by applying a developed version inspired by the game Ms. Pac-Man Maze Madness. In the VR world, challenges of this kind are usually different from the normal ones pertaining to the real world. For in a normal game, for instance, the player can observe the enemy moves on the screen, which is not the case in a VR game context. Besides, the player would be totally immersed in the game environment, so he/she could not follow the enemy moving, but he may hear its voice coming towards him as the sound level gets higher. Moreover, for the sake of making the player feel the VR interaction more intensely, we have reckoned it rather convenient to introduce other types of obstacles, such as moving walls to be pushed to open the gateway for the ball to move forward. In addition, and when the user pushes his avatar to the final destination level, it would automatically move on to the second level’s starting point together with the ball, overwhelmed with the feeling of being teleported from an environment to another. As simple graphics have been used in our first sketchy design of the game, whereby simplicity would make the user feel rather comfortable while getting immersed in a totally non-real environment, such a procedure is considered to help him/her get rid of confusion throughout the playing process.

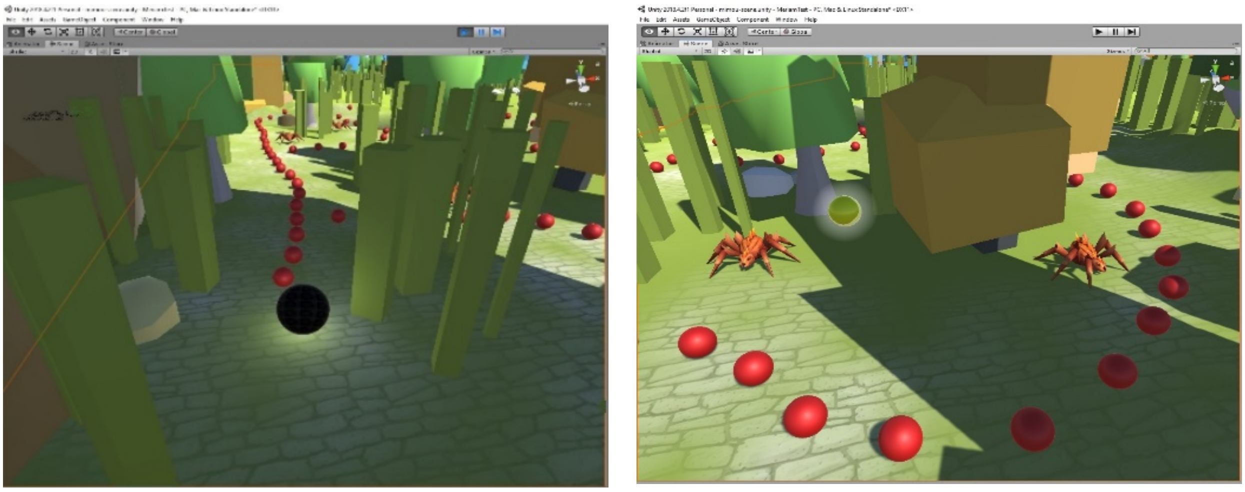


Figure 112. the XR game objects: a) game kick off b) game progress

Playability evaluation

Creating an XR environment with constructive ideas and stimulating goals is the fundamental part of the whole study. It is confirmed in our previous studies that visuals affect the satisfaction of the game. After evaluating the effect of visual components on the user’s satisfaction, we have applied some expert methods to confirm the affirmations of our creation. We chose the idea to evaluate its playability. Starting by testing the level of playability of the game based on Nielsen’s developed list of the ten evaluation heuristics, through literature reviews, game heuristics have been established by multiple individuals (Desurvire et al. , 2004). The “HEP,” as the Heuristics for Evaluating Playability, is one of these developments. A comprehensive set of heuristics to evaluate the game’s playability for the gamer (Nielsen, 1994). We have chosen one heuristic evaluation for each characteristic of the game’s playability factors. First, the gameplay; second, the game visuals; third, the mechanics and the game’s usability. We have arranged an evaluation session with three individuals working on the XR interface creation to explore the game's playability. The Playability of the game is a standard measurement system of the game's quality; it is the idea and the simplicity level of the Game playing story and its duration type. First, define whether the gameplay and the game's visual level are acceptable by answering the following questions: Does the game's idea appear to the user while playing it, and is the game enjoyable enough to replay? Also, how much does the Player show interest in the game's storyline idea? With the help of some XR experts in experimenting with the game and answering this question, it is convenient that the idea and the type of the chosen game are clear, straightforward, and enjoyable. The experience of the game's idea has an old relation to the user's real life and grabs their interest in an obvious way because it is based on two of the most famous arcade game in human history, Pac man and snake. Second, verifying the game's mechanics and usability to prove its efficiency, the questions are: are the game's score and goal identified easily for the player? With no doubt that the game's interactivity is relatively easy, and as it is based on famous games, its goals are clear for the user, as for the score and the level that the player reaches appear on the screen all the time to facilitate the information presented for him. is the menu well organized, minimalist, and experienced as part of the game? This act reveals if the game provides a good and a rich experience with fewer pointless distractions or difficulties that the designers did not intend. The game's Quality testing, functionality, and adaptability are necessary to evaluate the success level of the game. We conducted a deep discussion about the game elements and its idea, where the expert determined their playability level, as seen in Figure 113.

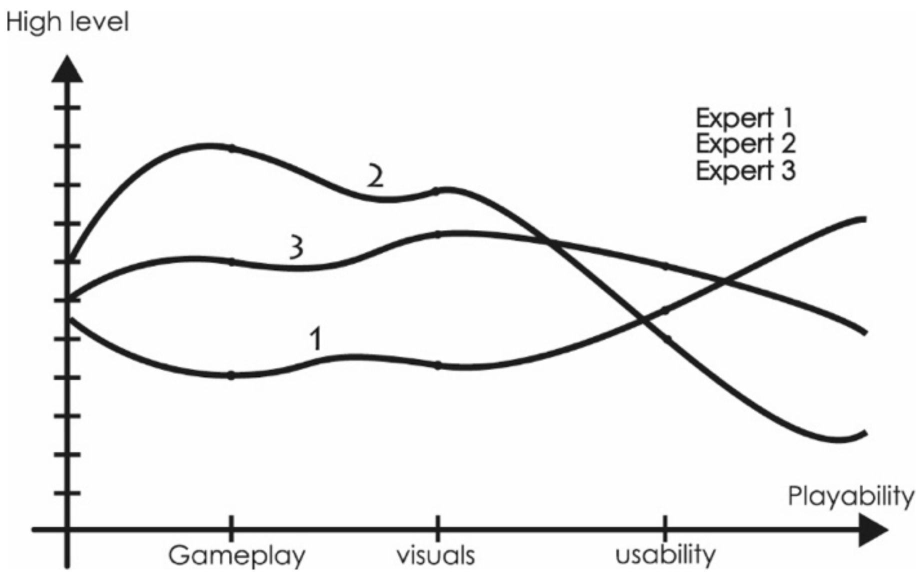


Figure 113. *Game quality testing*

Conclusion

The creation of games is spreading worldwide, involving different fields in it. It has become one of the most used fields with multiple purposes other than its real ones, such as education, healthcare, and even the military. In creating a VR environment, many purposes are followed; we find the impact of the graphical aspect in VR games where the quality of the graphics and the objects included in the game scenes are significant and contain several characteristics.

- The feel of graphics affects the playability; also, the fictional sensation becomes real when it affects two of the five senses: the haptics and the visual.

- Playability (P) consists mainly of gameplay, mechanics, and graphics. Playability is the quality or state of being playable, the quality of being easy and enjoyable to play, or the degree to which a video game is. It is the level of satisfaction (S) that the player reaches. (P) = (S)

References

Cheah, I., Shimul, A. S., & Phau, I. 2022. Motivations of playing digital games: A review and research agenda. Psychology & Marketing, 39(5), 937–950, https://doi.org/10.1002/mar.21631.

Desurvire, H., Caplan, M., & Toth, J. A. 2004. Using heuristics to evaluate the playability of games. In CHI'04 extended abstracts on Human factors in computing systems (pp. 1509-1512), https://doi.org/10.1145/985921.986102.

Egenfeldt-Nielsen, S., Smith, J. H., & Tosca, S. P. 2013. Understanding video games: The essential introduction, ISBN-1138849820 Routledge.

Lrez, B. M. 2014. Graphics, playability, and social interaction, the greatest motivations for playing Call of Duty. Educational reflections. Journal of New Approaches in Educational Research (NAER Journal), 3(1), 34–41. https://DOI:10.7821/naer.3.1.34-41.

Nielsen, J. 1994. Usability inspection methods. In Conference companion on Human factors in computing systems (pp. 413-414), https://doi.org/10.1145/259963.260531.

Sánchez, J. L. G., Simarro, F. M., Zea, N. P., & Vela, F. L. G. 2009. Playability as Extension of Quality in Use in Video Games. In international conference in computer science.

Alterations in Alternating Attention Following Short-Term Virtual Reality Training Among Amateur E-Athletes

Maciej Lachowicz[1], Dariusz Jamro[2], Anna Serweta-Pawlik[3], Alina Żurek[4], and Grzegorz Żurek[1]

[1]*Department of Biostructure, Wroclaw University of Health and Sport Sciences, Wroclaw, Poland*
[2]*Department of Physical Education and Sport, General Tadeusz Kosciuszko Military University of Land Forces, Wroclaw, Poland,*
[3]*Department of Occupational Therapy, Wroclaw University of Health and Sport Sciences, Wroclaw, Poland.*
[4]*Institute of Psychology, University of Wrocław, Wrocław, Poland*

Corresponding author: Dariusz.Jamro@awl.edu.pl, Maciej.lachowicz@awf.wroc.pl

Keywords: E-Sports, Virtual Reality, Alternating Attention, Cognitive Skills

Abstract

In the competitive world of e-sports, cognitive skills like alternating attention are crucial. We conducted a study on 64 amateur e-athletes to explore the effects of short-term virtual reality (VR) training on these skills. Participants were divided into two groups: one receiving VR training and the other as a control group. We used the Color Trail Test to measure alternating attention before, after 8 days of VR training, and 31 days later as follow-up tests. Results showed significant improvements in the VR training group, highlighting the potential of VR for enhancing e-athlete cognitive functions. This study suggests that tailored training programs can give e-athletes a competitive edge. While this research has promising implications, it's important to note potential compliance issues and reliance on self-report measures. Further studies on executive functions in e-sports are needed for a comprehensive understanding of these effects.

Introduction

In the dynamic and rapidly evolving realm of e-sports, where players engage in intense competitive gaming across various genres and platforms, the significance of cognitive abilities, including alternating attention, cannot be overstated (Ding et al., 2018). E-sports competitions often involve complex strategic gameplay, demanding high levels of cognitive engagement, rapid decision-making, and precise motor responses. In high-stakes tournaments and team-based competitions, e-athletes must synergistically collaborate with their teammates, communicate effectively, and synchronize their actions. High alternating attention among team members contributes to cohesive teamwork, synchronized decision-making, and strategic coordination (Bertran & Chamarro, 2016). By optimizing these cognitive skills, e-athletes can elevate their individual performance while bolstering the overall synergy and effectiveness of their team. Moreover, the multidimensionality of e-sports necessitates a diverse set of cognitive skills depending on the game genre. Each game genre places unique cognitive demands on e-athletes, making it essential to develop and hone a range of cognitive abilities tailored to the specific gameplay requirements. One of such crucial cognitive functions is alternating attention (AA), which encompasses the capability to flexibly shift focus between multiple tasks, stimuli, or game objectives (Worringer et al., 2019). Effective task-switching and divided attention enable e-athletes to juggle multiple gameplay aspects, adapt to evolving situations, and allocate their cognitive resources efficiently. While previous studies have explored the benefits of VR training in cognitive domains, such as memory (Optale et al., 2010; Brooks et al., 2003), decision-making (Pagé et al., 2019), spatial awareness (Rasheed et al., 2015), visuospatial memory (Serweta-Pawlik et al., 2023) and other cognitive functions, to our knowledge, there have been no studies to date regarding VR training of alternating attention in amateur e-athletes. Therefore, this study aims to bridge this gap by investigating the impact of short-term VR training on this cognitive function. By exploring the potential of VR training, we aim to provide valuable insights into the optimization of cognitive training strategies for e-sports.

Materials and Methods

Participants

A total of 64 participants, comprising 42 men and 22 women with a mean age of 22.7 ± 0.66, were recruited as subjects for this study. A pre-study questionnaire was administered to gather pertinent information regarding the types of games most frequently played, average daily gaming time, e-sports experience, and age. No statistically significant differences were observed between the E and C groups concerning average daily gaming time, e-sports experience, age, or initial levels of alternating attention. Using a simple 1:1 randomization method, the participants were assigned to either the experimental group (E) consisting of 32 individuals who underwent training for eight consecutive weekdays, or the control group (C) comprising 34 individuals. The study experienced a dropout rate of 21.22%, only participants who completed the pre, post, and follow-up assessments were included in the final analysis. The well-being of participants throughout the VR intervention was closely monitored using the Virtual Reality Sickness Questionnaire, administered prior to and following each training session (Kim et al., 2018).

Measurements

In the present study, assessments of alternating attention were conducted before the first training session and 30 minutes after completing the last training session for the E group. For the C group the pre- and post-tests were performed with an 8-day interval between assessments. Additionally, follow-up tests were conducted for both groups 31 days after the post-tests to evaluate any potential long-term effects of the training interventions. The study utilized the Color Trail Test (CTT) to assess the focus of attention and executive functions. The test encompasses two inseparable components: CTT-1 and CTT-2, administered sequentially. CTT-1 measures visual search (VS), sustained attention (SA), and graphomotor ability. On the other hand, CTT-2 provides supplementary insights into the subject's capability for attentional divisibility, alternating attention (AA) and sequential information processing. During the task, participants are required to meticulously connect the numerals in ascending order, adhering to the prescribed alternating color pattern and ensuring an uninterrupted sequence without lifting the pencil from the paper. Subsequently, the accuracy and duration of this task are methodically recorded (CTT Professional Manual, PAR, accessed on 3rd september 2023). The primary measure is the time, measured in seconds, to complete the actual task in each part.

Study design

This was a two-group, randomized, single blind, controlled trial comparing the VR training group to control group. E group underwent eight immersive VR training sessions utilizing the game Beat Saber. These training sessions, with a duration of approximately 15 minutes each, were conducted using the Valve Index VR headset. Beat Saber is a VR game that involves wielding virtual lightsabers to slice through colorful blocks synchronized to the rhythm of the song. The game's difficulty level gradually increased over time, progressing from the normal level to the expert level every two days.

Statistical analysis

Statistical analyses were conducted using Statistica v. 13 software, provided by StatSoft Poland, within the Department of Biostructure at the Wroclaw University of Health and Sport Sciences, which adheres to ISO 9001 certification for maintaining high-quality standards. The statistical significance level was set at $p \leq 0.05$ to determine the presence of statistically significant findings. The normality of variables' distribution was assessed using the Shapiro-Wilk test. To examine differences within groups across the pre, post, and follow-up tests, Analysis of Variance (ANOVA) for repeated measures was employed, post-hoc Honestly Significant Difference (HSD) Tukey's tests were conducted to perform specific pairwise comparisons.

Results

The Analysis of Variance (ANOVA) for repeated measures regarding executive functions measured in pre, post, and follow-up tests showed statistically significant differences. Specifically, within the E group, enhancements were evident in both CTT tests. Subsequent post-hoc HSD Tukey testing, further substantiated the observed improvements indicated by the ANOVA analyses in the pre vs. post and pre vs follow-up comparison, thereby indicating a sustained training effect resulting from the interventions administered. Contrarily, the C group did not exhibit any statistically significant changes in the measured executive functions. Additional comparison regarding post and follow-up tests between group E and C showed statistically significant difference.

Table 7. CTT results as mean values ± standard deviation.

Group	Pre Test		Post Test		Follow-Up Test	
	CTT-1	CTT-2	CTT-1	CTT-2	CTT-1	CTT-2
Experimental	33.51 ± 3.6	63.54 ±5.41	24.21 ± 2.15	50.54 ± 2.99	25.23 ± 2.31	51.31 ± 4.36
Control	32.31 ± 3.7	61.36±4.7	29.37 ± 2.53	59.86 ± 3.61	30.76 ± 2.55	59.32 ± 3.33

Table 8. Results of ANOVA for repeated measures and HSD Tukey test

ANOVA for repeated measures			HSD Tukey Test p-values		
Test	Group	p-values	Pre vs Post	Post vs Follow-Up	Pre vs Follow-Up
CTT-1	E	0.0001	0.0001	0.8644	0.0002
	C	0.0555	0.3609	0.7495	0.7952
CTT-2	E	0,0001	0.0003	0.9671	0.0006
	C	0.5377	0.8596	0.7555	0.9802

Discussion

This study aimed to investigate the effectiveness of VR training in improving alternating attention. Our findings add alternating attention to the repertoire of trainable cognitive functions within the domain of VR-based interventions. When considering e-sports in the context of team-based activities, albeit the limited body of research specifically addressing executive functions within e-sports, pertinent insights can be drawn from a study conducted by Da Waelle et al. (2021). This study examined a group of children engaged in team sports and revealed that such collective endeavors not only demand but also foster the development of advanced executive functions, surpassing those required by individual-paced sports. Intriguingly, a differential pattern emerged for children participating in self-paced sports, where no discernible advantage in executive functioning was observed compared to non-athletic counterparts. Considering the shared attributes of team-based dynamics and competitive nature between e-sports and traditional team sports, it can be inferred that the demand for heightened executive functions aligns in both domains. Consequently, the prospect of utilizing VR-based training emerges as a valuable resource not only for aspiring amateur e-athletes aiming for elevated performance standards but also for conventional athletes engaged in team-based sporting disciplines.

Conclusion

Our study findings align with previous research that has showcased the promising advantages of VR technology in augmenting cognitive capabilities (Optale et al., 2010; Pagé et al., 2019; Serweta-Pawlik et al., 2023). Specifically, the experimental group (E) displayed discernible enhancements in alternating attention metrics. These outcomes offer valuable insights into the capacity of customized training interventions to induce targeted enhancements in cognitive proficiencies. The observed progress in alternating attention underscores the potential value of these interventions as efficacious tools for the preparatory stages of e-sports engagement.

Limitations

Despite our rigorous efforts to ensure a comprehensive study design, it is important to acknowledge several inherent limitations in this investigation. One notable limitation is the lack of direct supervision over participants' activities during weekends. Another limitation pertains to the reliance on self-report measures, particularly the questionnaires employed to gather pertinent data. Self-report instruments are susceptible to various response biases, encompassing social desirability bias and recall biases.

References

Bertran, E., & Chamarro, A. 2016. Videogamers of League of Legends: The role of passion in abusive use and in performance. Adicciones, 28(1). https://doi.org/10.20882/adicciones.787

Brooks, B. M., & Rose, F. D. 2003. The use of virtual reality in memory rehabilitation: current findings and future directions. NeuroRehabilitation, 18(2), 147-157. https://doi.org/10.3233/NRE-2003-18207

Color Trails Test (CTT) Professional Manual|PAR. Available online: https://www.parinc.com/Products/Pkey/77 (accessed on 3 september 2023).

Ding, Y., Hu, X., Li, J., Ye, J., Wang, F., & Zhang, D. 2018. What makes a champion: the behavioral and neural correlates of expertise in multiplayer online battle arena games. International Journal of Human–Computer Interaction, 34(8), 682-694. https://doi.org/10.1080/10447318.2018.1461761

Kim, H. K., Park, J., Choi, Y., & Choe, M. 2018. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. Applied ergonomics, 69, 66-73. https://doi.org/10.1016/j.apergo.2017.12.016

Optale, G., Urgesi, C., Busato, V., Marin, S., Piron, L., Priftis, K., ... & Bordin, A. 2010. Controlling memory impairment in elderly adults using virtual reality memory training: a randomized controlled pilot study. Neurorehabilitation and neural repair, 24(4), 348-357. https://doi.org/10.1177/1545968309353328

Pagé, C., Bernier, P. M., & Trempe, M. 2019. Using video simulations and virtual reality to improve decision-making skills in basketball. Journal of sports sciences, 37(21), 2403-2410. https://doi.org/10.1080/02640414.2019.1638193

Rasheed, F., Onkar, P., & Narula, M. 2015. Immersive virtual reality to enhance the spatial awareness of students. In Proceedings of the 7th Indian Conference on Human-Computer Interaction (pp. 154-160). https://doi.org/10.1145/2835966.2836288

Serweta-Pawlik, A., Lachowicz, M., Żurek, A., Rosen, B., & Żurek, G. 2023. Evaluating the Effectiveness of Visuospatial Memory Stimulation Using Virtual Reality in Head and Neck Cancer Patients—Pilot Study. Cancers, 15(6), 1639. https://doi.org/10.3390/cancers15061639

Worringer, B., Langner, R., Koch, I., Eickhoff, S. B., Eickhoff, C. R., & Binkofski, F. C. 2019. Common and distinct neural correlates of dual-tasking and task-switching: a meta-analytic review and a neuro-cognitive processing model of human multitasking. Brain Structure and Function, 224, 1845-1869. https://doi.org/10.1007/s00429-019-01870-4

Designing an experimental virtual museum for extended social inclusion through multimodality

Eleftherios Anastasovitis, Georgia Georgiou, Eleni Matinopoulou,
Spiros Nikolopoulos, and Ioannis Kompatsiaris

*Multimedia Knowledge and Social Media Analytics Laboratory (MK-Lab),
Information Technologies Institute (ITI),
Centre for Research and Technology Hellas (CERTH)*

Corresponding author: anastasovitis@iti.gr

Keywords: Virtual Reality, Experimental Virtual Museum, Cultural Heritage, Inclusion, Multimodality

Introduction

Immersive and emerging technologies have met an impressive evolution in the last years and their presence is in everyday life. To this direction helped the wide use of the democratized game engines that play the role of the platforms that combine heterogeneous technologies. Nowadays, the virtual museum is a type of application that is used for communication and understanding, mainly in the sector of cultural heritage.

The main objective of the *MuseIT* research project is the promotion of inclusion, participation, accessibility and equal opportunities for all through the development of a platform for remote immersive co-creative engagement with cultural assets and experiences. The ongoing *MuseIT* project aims to improve: (a) accessibility of cultural assets by development of multisensory representations and alternative expressions to enable engagement by the public regardless of functional or sensory impairments, (b) engagement with cultural assets and cultural co-creation by a broader public (regardless of variations in abilities and perceptual modalities), and (c) methodologies for preservation and safeguarding of cultural heritage with inclusion at its core.

This contribution presents the initial state of the production phase for the virtual exhibition about the immersive cultural heritage experiences. Various research teams for further multimodal and multisensory experiments through haptics and emerging technologies will deliver this virtual space. This poster depicts the technical aspects of the virtual museum that will be used as an experimental research space.

Key characteristics of the experimental virtual museum

As a first step, the research team searched and retrieved digital representations of cultural heritage assets by open platforms and aggregators. Through an extensive selection we selected a representative group of assets in different modes, such as 2D and 3D representations. As a next step, we applied the recommendations of virtual museology to design the thematic areas of the virtual museum (Table 9).

The refinement of the 3D models, as well as any further 3Dcontent creation will be made through *Cinema4D*. The Unity3D game-engine will be used for the development of the virtual museum, according to the blueprints of the virtual museology plan. Based on gained expertise from previous projects (Anastasovitis et al., 2017; Anastasovitis & Roumeliotis, 2018; (Anastasovitis et al., 2020; Anastasovitis & Roumeliotis, 2021; Anastasovitis et al., 2021; Anastasovitis et al., 2022), the research team defined a scenario regarding the VR application. The virtual collections will be assisted by different modalities, such as text and audio. The VR headset will be *Oculus Quest Pro* and *Oculus Quest 3*.

To resolve issues regarding the accessibility, the text will be adaptive and its oral description will be always available to any user. Moreover, ambient sounds for the whole virtual space, as well as specific sounds in each exhibit will also be present. In addition, the lighting of the room will be also adaptive (color, intensity, position), for the purposes of the forthcoming experiences.

In collaboration with the research group that is responsible for the electroencephalogram (EEG) experiments, we will set up a collection of triggered colliders that will help in the analysis of the behavior of the subjects. For example, when the user will reach in front of an artefact (trigger collider A) and examines it (trigger through ray-casting), the EEG researchers will retrieve useful game analytics that will be combined with the electrical signals.

Table 9. Knowledge transfer for each thematic area of the experimental virtual museum.

Thematic	Knowledge transfer	Means of communication
The city landscape	Aspects of a city's transformation through the time	Maps, city plans, VR experiential
The city's history	Aspects of a city's historical milestones	Artworks, VR experiential
People's religion	The religion as a key unifying factor and representative aspects of people's identity	VR experiential
The people	Important personalities of the city's history that operate as reference points for the people	Artworks, VR experiential

Conclusion and future work

In this contribution the design of an immersive virtual museum that acts as an experimental digital space where the aspects of multimodality and social inclusion will be examined, was presented. The use of virtual reality and the creative industry of videogame for designing a virtual museum for research purpose is the core of this presentation. The virtual museum is intended to be used for experiments with EEG and haptics. The analysis of the results is expected to provide evidences and recommendations for extended reality museums in the era of Metaverse.

Acknowledgements

The *MuseIT* project "*Multisensory, User-centered, Shared cultural Experiences through Interactive Technologies*" is co-funded by the European Union under the Horizon Europe Framework Program and the grant agreement N°101061441.

References

Anastasovitis, E., Ververidis, D., Nikolopoulos, S., & Kompatsiaris, I. 2017. Digiart: Building new 3D cultural heritage worlds. In 2017 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 7-9 June 2017, Copenhagen, Denmark, (pp. 1-4). IEEE. https://doi.org/10.1109/3dtv.2017.8280406.

Anastasovitis, E., & Roumeliotis, M. 2018. Virtual Museum for the Antikythera Mechanism: Designing an immersive cultural exhibition. In 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Munich, Germany, 16-20 Oct. 2018, pp. 310-313, https://doi.org/10.1109/ismar-adjunct.2018.00092. IEEE.

Anastasovitis, E., Georgiou, G., Nikolopoulos, S., & Kompatsiaris, I. 2020. EPANASTASIS-1821: Reviving the naval history of the Revolution through full-immersive Virtual Reality. In Helin, K., de Antonio, A., & Reyes-Lecuona, A. (Eds.) (2020). EuroVR 2020 Application, Exhibition & Demo Track: Proceedings of the Virtual EuroVR Conference. VTT Technical Research Centre of Finland. VTT Technology No. 381, pp.65-70. https://doi.org/10.32040/2242-122X.2020.T381, Valencia, Spain, 25-27 November 2020.

Anastasovitis, E. & Roumeliotis, M. 2021. Creative Industries and Immersive Technologies for Training, Understanding and Communication in Cultural Heritage. In: Ioannides M., Fink E., Cantoni L., Champion E. (eds) Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection. EuroMed 2020. Lecture Notes in Computer Science, vol 12642, pp.450-461, Cyprus, 02-05 November 2020. Springer, Cham. https://doi.org/10.1007/978-3-030-73043-7_37

Anastasovitis, E., Tsaknaki, E., Georgiou, G., Nikolopoulos, S., & Kompatsiaris, I. 2022. ECHOES from the past through immersive technologies and creative industry (Best Application Poster Award). In Helin, K., Reyes-Lecuona, A., & Runde, C. (Eds.) (2022). The application track, posters and demos of EuroXR: Proceedings of the 19th Annual EuroXR Conference - 2022. VTT Technical Research Centre of Finland. VTT Technology, No. 408, pp.91-94. https://doi.org/10.32040/2242-122X.2022.T408 Stuttgart, Germany, 14-16 September 2022.

Georgiou, G., Anastasovitis, E., Nikolopoulos, S., Kompatsiaris, I. 2021. EPANASTASIS-1821: Designing an Immersive Virtual Museum for the Revival of Historical Events of the Greek Revolution. In: Ioannides M., Fink E., Cantoni L., Champion E. (eds) Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection. EuroMed 2020. Lecture Notes in Computer Science, vol 12642, pp.334-345, Cyprus, 02-05 November 2020. Springer, Cham. https://doi.org/10.1007/978-3-030-73043-7_28

Demos

Advancing Collaborative Remote Operations through Social Extended Reality: A Case Study in Search and Rescue

Tessa Klunder, Bas Binnerts, Galit Rahim, Sylvie Dijkstra-Soudarissanane, and Omar Niamut

TNO, The Hague, the Netherlands

Corresponding author: Tessa.klunder@tno.nl

Keywords: Search and Rescue, Social-XR, Remote Expertise, Augmented Reality, Virtual Reality, Mixed Reality, XR-Collaboration

Introduction

Existing 2D video communication tools such as Microsoft Teams or Zoom have proven valuable remote interactions. However, the next step in this technological development is to replicate the feeling of social and spatial presence, by allowing people to experience being together. Drawing from cutting-edge Virtual Reality (VR) techniques and knowledge about remote collaboration within our research program Social Extended Reality (SXR) (TNO, 2023a), a virtual collaboration space has been conceived. Through SXR technology, people are united in a common virtual environment, facilitating data sharing enabling face-to-face interactions and decision-making irrespective of the geographic location.

This innovative virtual collaboration space holds potential in supporting complex and time-critical operations involving multiple stakeholders, such as in the context of Search and Rescue (SAR) missions. The need for novel forms of communication is clear as in 2015, as the result of a diving accident in the North Sea, the Dutch Safety Board investigated the handling and communication between the coastguard and other stakeholders during the incident (Dutch Safety Board, 2016). The Board concluded that "in the absence of any central coordination, the participating parties took independent decisions and, without the consent of the Coastguard Centre, independent action". SAR operations are characterized by processing and sharing a lot of information often under immense time constraints, with mobile phones serving as the primary communication tools. The novel implemented virtual collaboration space not only provides team members with a shared situational overview of maritime operation, but also enhances collaboration through 3D representations of team members enabling better interpersonal and social interaction. This paper outlines the configurations, as well as the feedback from the Dutch Search and Rescue Team on the demonstrator.

Virtual Collaboration Space: Multi-modal Experience Design

The proposed technical architecture, depicted in Figure 114, empowers three geographically dispersed users with a 3D virtual collaboration space. This collaborative system is designed to facilitate remote users that will join a common space with diverse modalities during a Search and Rescue (SAR) mission. Two users that are captured with RGBD cameras join as 3D point clouds and interact with a third user that is using a tablet. All streams are sent directly between users to enable low latency and real-time interaction amongst the users as can be seen in the video (TNO, 2023b) created about this demo.

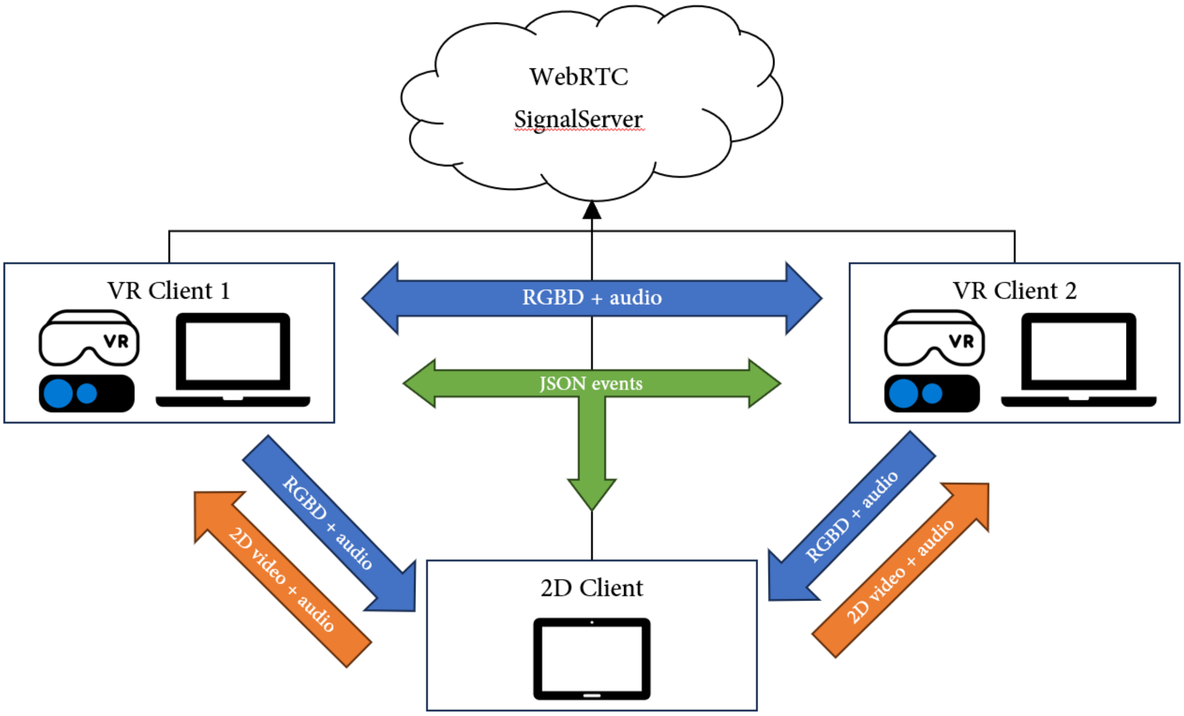


Figure 114. System architecture using WebRTC for peer-to-peer streaming.

User capture

The VR users are captured with RGBD cameras, concatenating the RGB and depth images as a single stream to the other clients (Gunkel et al., 2023). The captured resolution is 512x512, which becomes 1024x512 when combining RGB and depth information, and contains a depth range of 1.5 meters at 30 frames per second. At the receiver side, these images are used to reconstruct life-size and real-time photo-realistic point clouds in the virtual collaboration space. Combined with spatial audio, the VR users get a sense of being virtually present with the other users. Changes made to the environment, such as switching data layers on a nautical chart or pointing at a specific location, are also shared in real-time between all clients. Enabling users to perceive each other's actions and manipulations within the shared virtual space which improves collaboration between users. To enable data synchronicity, one client acts as a controller and has priority when adding changes to the scene. Pointer locations are sent regardless of role and streamed directly to all clients.

Virtual Reality has many benefits, but also limitations when it comes to remote or difficult to reach locations. To mitigate this constraint for system users, a 2D interface is incorporated into the session accessible via a tablet or laptop, requiring only a stable internet connection for joining. The user is captured using the 2D integrated camera of the device and uses only the RGB images from the VR users without having to process the depth images. The 2D user sees a Teams- or Zoom-like grid of webcam images from the other users and a top-down view of the virtual scene. On the other end, the VR users will receive the 2D camera stream as a floating TV screen in their three-dimensional space. Creating a multi-modal interaction (VR, 2D) system, without the limitations of hardware or the physical environment.

Networking

To connect multiple users seamlessly and independently from their geographic location a WebRTC signaling server (Kutza, 2023) is set up. This server ensures a direct peer-to-peer connection, optimized for low-latency RGBD streaming between all users, while simultaneously providing the possibility to stream 2D video using the same connection. Additionally, all user events and pointer locations are sent as JSON through the WebRTC connection.

Demo set-up

Our demonstrator is an embodiment of the technical architecture described in section 2. We use a typical SAR mission scenario for context and functional requirements. The SAR team is formed by three different actors: the Mission Coordinator (MC), the Drone Operator (DO), and the On-Scene Coordinator (OSC). The scenario was tested by incorporating all the technical modules into a fully integrated demonstrator. Figure depicts the VR view of the virtual collaboration room.

Scenario

A report of a missing person is received by the Coast Guard. The Mission Coordinator and Drone Operator take place in the virtual collaboration room and the local On-Scene Coordinator is called in virtually. Relevant information is integrated on the interactive 'nautical chart table' where the MC can switch between different information layers. The layers contain information about (i) the environmental conditions at sea, (ii) the search area based on simulated drift results for a drowning person, (iii) the selected search patterns for the different assets at sea, and (iv) the locations of the search assets during the operation.

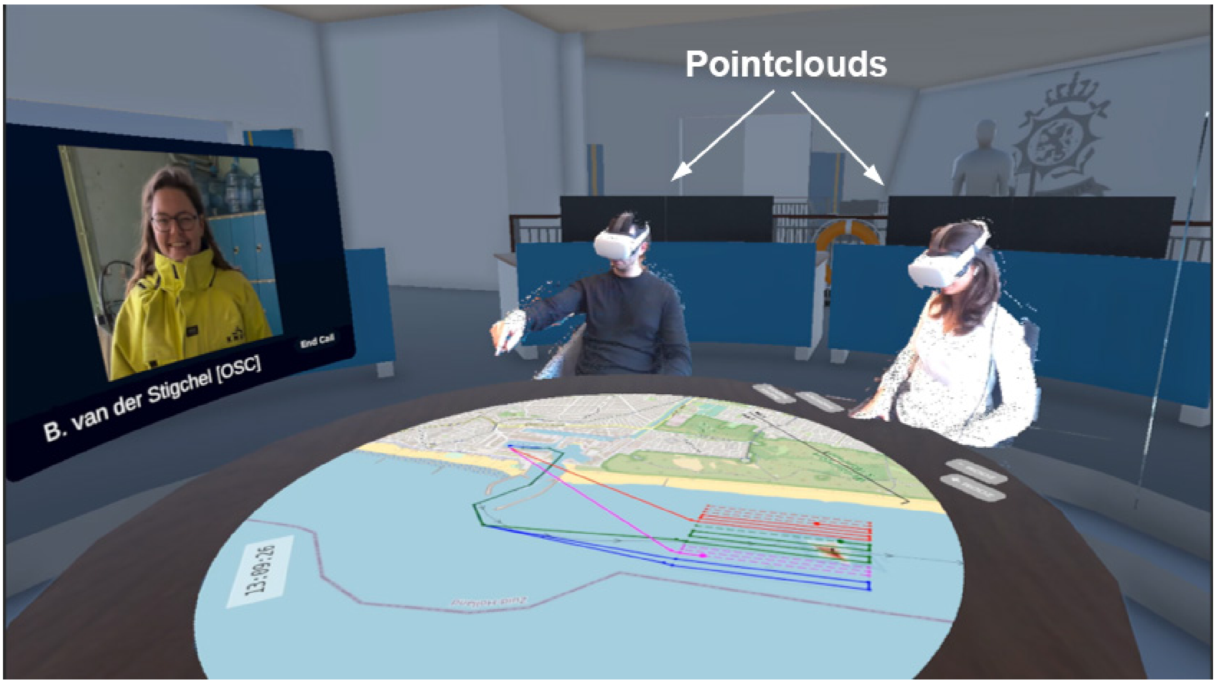


Figure 115. *Two VR and one 2D user inside the virtual collaboration space.*

With all information in one place, the team can construct the best plan together and monitor the operation at a distance. The MC makes an initial proposal, and the plan can be adjusted and confirmed in the system based on local knowledge of the OSC or the estimates of the DO. Once on site, the units begin executing the search plan, while their location and sailing route are visible on the interactive map.

This joint picture makes the plan clear to everyone and allows for adjustments in the event of disruptions. Taking over tasks is also easier, so that the operation can always continue.

Physical setup

The demo set-up consists of two clients that use VR-headsets (i.e. Meta Quest 2) connected by cable to a VR compatible laptop. An RGBD camera (i.e. Microsoft Azure Kinect) is connected to the laptop, which via a wireless router enables the peer-to-peer connection between the clients. A third client can access the virtual collaboration space using a tablet.

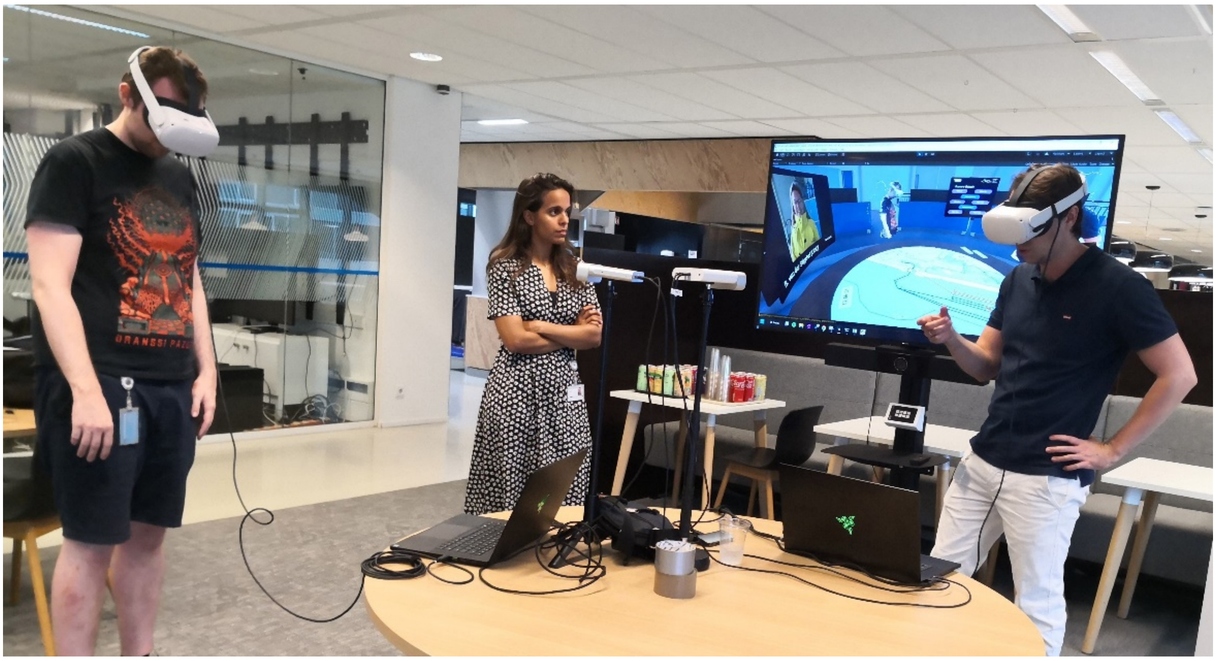


Figure 116. The VR setup for two users in the Virtual Collaboration Space.

Main Results

We presented and discussed this demonstrator to main stakeholders such as the Dutch Coastguard and the Royal Netherlands Sea Rescue Institution (KNRM). These stakeholders highly valued key aspects. For example, the integration of various detailed data on the map emerged as a particularly advantageous feature combined with the intuitive interaction design and user-friendly interface. The fact that users require minimal guidance regarding controls, interaction, and application purposes underscores the effectiveness of the design in facilitating seamless user engagement. The Dutch Coastguard's endorsement of this integration as a potential source of significant efficiency gains in both preparation and decision-making processes further solidifies the application's practical utility. The KNRM singled out the 2D tablet version of the system for its exceptional value as the 2D tablet version not only streamlined the user’s involvement in discussions while at sea but also eliminated the need for complex systems or head-mounted devices, enhancing operational convenience.

However, it is important to acknowledge certain limitations of the current version. The reliance on static data for creating the nautical chart means that the application's potential is currently demonstrated as a proof-of-concept rather than a fully integrated and flexible solution. Transitioning towards seamless integration would demand considerable efforts to harmonize various components and ensure timely availability of data as the real-time aspect of this type of communication is essential. Additionally, the issue of limited bandwidth, particularly in maritime settings, cannot be ignored. Given the constraints posed by the scarcity of bandwidth at sea, careful consideration must be given to optimize data transmission and maintain the highly interactive functionality.

Conclusion and Future Work

The Social-XR Virtual Collaboration Space provides a new method for complex and time-critical operations like Search and Rescue (SAR) missions. Real-time photo realistic capture, shared object interactions inside a common space and integrated information on an interactive map demonstrate the possibilities of future collaborations. Tests with the stakeholders resulted in positive feedback and the endorsement of a potential source of significant efficiency gains.

Effort is still needed in terms of overcoming connectivity challenges at sea, as a stable and robust network is required for the collaboration space to reach its full potential. We plan on optimizing data transmission, stream stability and capture quality to ensure the important interactive aspect. Additionally, the application must be connected to the existing systems used by the Coast Guard and KNRM for seamless data integration. Further steps involve our development of a near cost drift model that can be integrated as an additional planning tool in the virtual space. Finally, future developments should focus on testing with stakeholders and eventually integrating with real SAR-missions.

References

Dutch Safety Board 2016. Rapport Summary Investigation of a diving accident and the provision of medical assistance on the North Sea, https://www.onderzoeksraad.nl/nl/page/4041/medische-hulpverlening-op-de-noordzee

Gunkel, S. N., Dijkstra-Soudarissanane, S., Stokking, H. M., & Niamut, O. A. 2023. From 2D to 3D video conferencing: Modular RGB-D capture and reconstruction for interactive natural user representations in immersive Extended Reality (XR) communication. Frontiers in Signal Processing, 3, 1139897. https://doi.org/10.3389/frsip.2023.1139897

Kutza, C. 2023. WebRTC Video chat, https://because-why-not.com/

TNO 2023 a. How Social XR (extended reality) and the metaverse reduce distances, https://www.tno.nl/en/digital/digital-innovations/digital-infrastructures/social-xr-extended-reality-distances/

TNO 2023b. Social XR for Search and Rescue. https://www.youtube.com/watch?v=XHVTS-TE7RQ

Unified Interface for Programming and Control of Industrial Robots

Yevhen Bondarenko, Simone Luca Pizzagalli, Vladimir Kuts, and Tauno Otto

Tallinn University of Technology, Department of Mechanical and Industrial Engineering

Corresponding author: yevhen.bondarenko@taltech.ee

Keywords: Industrial Robotics, Robot Operating System, ROS, Human-Machine Interaction

Introduction

Industrial robot systems traditionally use proprietary programming environments and languages, which differ depending on the system's manufacturer. This limits the potential for experimentation and development of new programs by researchers, enthusiasts and small businesses in the robotics field. This work presents a system based on the Robot Operating System (ROS)[22] and Unity[23], which allows for the control of any industrial robot from an interactive Unity environment. The system aims to be easy to set up while remaining hardware- agnostic by relying on existing standards from ROS community. Our approach provides a unified interface for industrial robot programming and creation of new applications using Unity's 3D visualization capabilities. We demonstrate the feasibility of our implementation by successfully controlling different types of industrial robots through the proposed system.

Architecture description

The proposed system's architecture is aimed at being modular to accommodate different robot models that can be connected to the Unity-based control interface. For this reason, the system contains two main parts: ROS module responsible for the simulation of the robot's digital twin (DT), and Unity module which provides the universal user interface (UI) for the operator to control and program the connected robot.

ROS Module Architecture

The system's ROS module is based on the MoveIt motion planning framework (Coleman et al., 2014) and ROS- Industrial[24] standard. MoveIt provides advanced functionality for motion planning and collision avoidance which can be used to create a DT of the robot and control it. ROS-Industrial standard provides restrictions for the ROS software packages aimed at supporting industrial robots, which makes the process of integration of new robot models straightforward and allows many setup steps to be automated. For instance, each official ROS-Industrial robot support package contains files describing the

[22] About Robot Operating System (accessed on 30-05-2023) https://www.ros.org/about-ros/
[23] Unity Real-Time Development Platform (accessed on 30-05-2023) https://unity.com/
[24] ROS-Industrial (accessed on 30-05-2023) https://rosindustrial.org/

kinematic chain of the corresponding robot and provides drivers for interfacing with the physical robot hardware from the ROS server. ROS-Industrial uses MoveIt node for simulation but provides a unified control interface in the form of ROS topics and services independent from the robot model used, as can be seen in the ROS module architecture diagram in Figure 117.

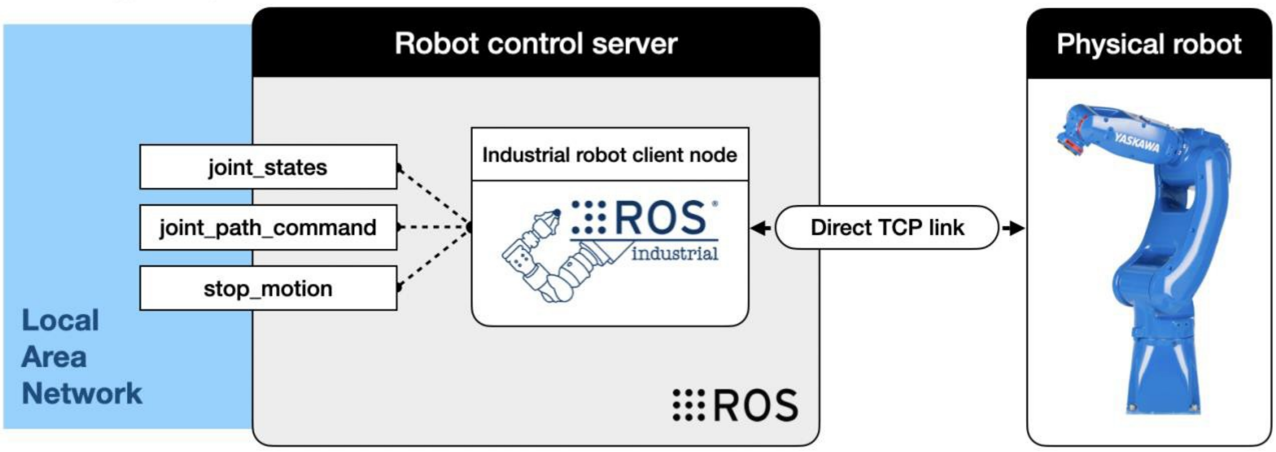


Figure 117. ROS side of the system.

Unity Module Architecture

Unity application for controlling the robot communicates with the ROS module by the means of ROS TCP Connector[25], a software component which allows to communicate to ROS servers from C# code used in Unity. We also use Unity URDF importer[26] project, which allows us to import 3D models of the robots into Unity directly from their ROS-Industrial packages, saving extra effort when setting up a new robot DT in Unity. The result is a virtual model of the robot which is linked through online connection to the MoveIt simulation in ROS, and can be controlled through C# scripts written in Unity. This allows us to build complex UIs and novel interaction methods for programming industrial robots using Unity's rich 3D toolset, which would not be available if we used only ROS. See Figure 118. For the architecture diagram of the Unity module.

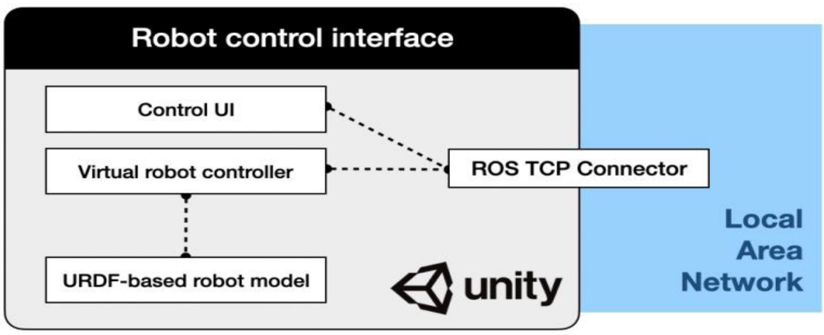


Figure 118. Unity side of the system.

Testing on Real Robots

To validate the developed system, it was used to configure the digital twins of three industrial robots available at TalTech IVAR Laboratory. The tested robot models include Motorman GP8, ABB IRB1600-6-12, ABB IRB 1200 and Omron TM5-900. A digital twin of each robot was created on the corresponding ROS server,

[25] ROS TCP Connector (accessed on 30-05-2023) https://github.com/Unity-Technologies/ROS-TCP-Connector
[26] ROS URDF Importer (accessed on 30-05-2023) https://github.com/Unity-Technologies/URDF-Importer

which can be connected to from Unity for control and experimentation. Two of the mentioned digital twins were successfully utilized in a user experience experiment conducted in TalTech IVAR Lab and Technical University of the Shannon. The experiment featured the use of a native teach pendant, Virtual and Augmented Reality interfaces to program the connected robots and execute the programmed trajectories on the physical hardware. The same experimental system will be presented as a demo accompanying this submission.

The native teach pendant for the ABB IRB 1600 and ABB IRB 1200 robotic arms is the model DSQC679 and the similar for both robots. Regarding the XR interface we decided to opt for full hand manipulation which is more natural and intuitive than controller-based interactions. All XR UI buttons and components can be interacted through an index finger poke gesture as in common physical hardware interface element. The headsets used for the AR and VR interfaces are Meta Quest2 and Hololens2 respectively.

The path planning tool provided in both AR and VR scenarios is presented as a sphere which has a few properties and components that facilitate the creation of the robot paths. The sphere itself is transparent, to avoid the occlusion of the robot and paths. Cartesian axes and origin point references are included in the tool and allow the correct positioning and orientation during the task. The center point is specifically meant to locate the newly created path points in the correct world space position. A UI button is provided next to the sphere, allowing for the rapid instancing of path points once the path has been initialized in the main interface. In both scenarios, the path planning tool can be grabbed through hand gestures and manipulation provided by both Oculus and MRTK SDKs for Unity. Figure 119 shows the AR UI setup for ABB IRB 1600, with the physical robot in the background, the AR setup and Path Planning Tool for ABB IRB 1200, and the VR UI for ABB IRB 1600.

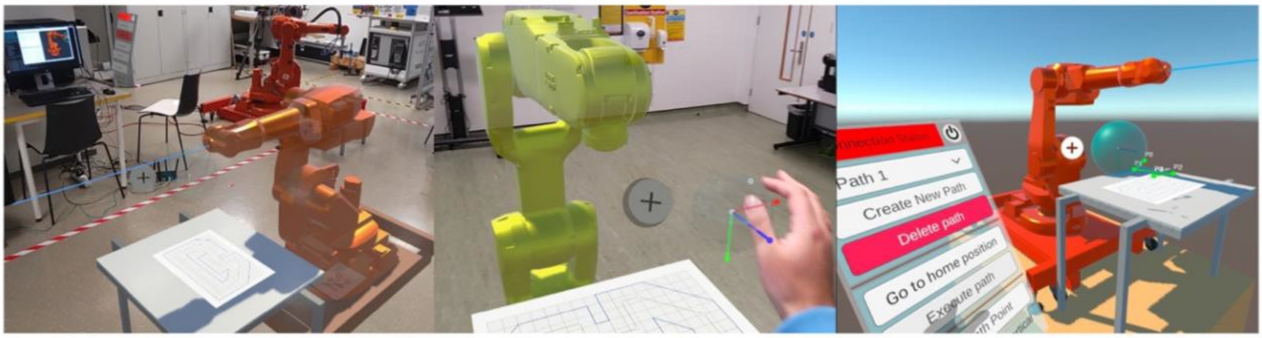


Figure 119. From the right the AR UI for ABB IRB 1600, the AR UI for ABB IRB 1200 and VR setup for ABB IRB 1600.

Other functions are provided in the main UI including the possibilities to create new path or select an existing one from the list, add path points to the existing path, reset the robot to the home position and execute the loaded paths. The paths are saved on the device in JSON format and loaded dynamically on request.

Conclusions

The developed system provides a streamlined way to create DTs of industrial robots compatible with a standardized interface in Unity engine, which can be used as a basis for the development of novel robot control scenarios in highly interactive 3D environments. The practicability of the system is already displayed by its use in the user experience experiments on real robots, with more potential developments to come in the future.

Reference

Coleman, D., Sucan, I., Chitta, S., & Correll, N. 2014. Reducing the Barrier to Entry of Complex Robotic Software: a MoveIt! Case Study, Journal of Software Engineering for Robotics, 5(1):3–16. https://doi.org/10.48550/arXiv.1404.3785

Author Index and Table of Figures & List of Tables

Author Index

Table of Figures

List of Tables

Series title and number

VTT Technology 422

Title	**EuroXR 2023** Proceedings of the 20th EuroXR International Conference
Author(s)	Kaj Hélin, Frédéric Noël, and Wolfgang Schäfer (eds.)
Abstract	The focus of EurXR 2023 is to present novel Virtual Reality (VR), Augmented Reality (AR) and Mixed Reality (MR) technologies, including software systems, display technology, interaction devices, and applications. Besides scientific papers reporting on new advances in the VR/AR/MR interaction technologies, the conference programme includes application-oriented presentations, creating a unique opportunity for participants to network, discuss, and share the latest innovations around commercial and research applications. As in previous years, we welcome industrial and academic exhibitors, as well as sponsors, all within the same exhibition area, to connect with our community. Our major priority is to provide authors the opportunity to prestigiously disseminate their innovative work within the wide community of end-users, from large scale industries to SMEs.
ISBN, ISSN, URN	ISBN 978-951-38-8785-8 ISSN-L 2242-1211 ISSN 2242-122X (Online) DOI: 10.32040/2242-122X.2023.T422
Date	November 2023
Language	English
Pages	196 p.
Name of the project	
Commissioned by	
Keywords	
Publisher	VTT Technical Research Centre of Finland Ltd P.O. Box 1000, FI-02044 VTT, Finland, Tel. 020 722 111, https://www.vttresearch.com

EuroXR 2023

Proceedings of the 20th EuroXR International Conference

The 20th EuroXR international Conference - EuroXR 2023 - taking place on 29 September - 1 December 2023 organized by the De Doelen in Rotterdam, The Netherlands.

The conference follows a series of succesfull European VR/AR conferences taking place since 2004 and is known as INTUITION. JVRC and recently EuroVR (Bremen 2014, Lecco 2015, Athens 2016, Laval 2017, London 2018, Tallinn 2019, Valencia 2020, Milan 2021 and Stuttgart 2022).

EuroXR 2023 will bring together people from reseach, industry, and commerce. Its members include technology developers, suppliers, and all those interested in Virtual Reality (VR), Mixed Reality (MR), including Augmented Virtuality (AV) and Augmented Reality (AR), and more globally 3D user interfaces, to exchange knowledge and share experiences, new results and applications, enjoy live demonstrations of current ans emerging technologies, and from collaborations for future work.

ISBN 978-951-38-8785-8
ISSN-L 2242-1211
ISSN 2242-122X (Online)
DOI: 10.32040/2242-122X.2023.T422

VTT

beyond the obvious