Library Declaration Form

University of Otago Library

Author's full name and year of birth: Wei Hong Lo,
(for cataloguing purposes) [redacted]

Title of thesis: ARSpectator — Enriching On-Site Sports Spectating with Augmented Reality

Degree: Doctor of Philosophy

Department: Department of Computer Science

Permanent Address: [redacted]

I agree that this thesis may be consulted for research and study purposes and that reasonable quotation may be made from it, provided that proper acknowledgement of its use is made.

I consent to this thesis being copied in part or in whole for

i) a library

ii) an individual

at the discretion of the University of Otago.

Signature: [redacted]

Date: 3/11/2022

ARSpectator — Enriching On-Site Sports Spectating with Augmented Reality

Wei Hong Lo

a thesis submitted for the degree of

Doctor of Philosophy

at the University of Otago, Dunedin,

New Zealand.

27 June 2022

This thesis is dedicated to my mum.

Abstract

Recent technological advancements in sports broadcasting provide an enhanced experience for broadcast viewers through visualizations, statistics, commentary, and better viewpoints. Unfortunately, on-site spectators often do not have the same access to such information. In this thesis, we introduce a novel system *ARSpectator*; an Augmented Reality (AR) approach that integrates event-related information into the on-site spectators' field of view. This thesis describes the overall system of *ARSpectator* but ultimately focuses on the research's visualization, interface and user experience aspects.

The thesis starts by describing the components of *ARSpectator* and their interactions. Due to limited on-site accessibility, we developed a prototyping framework that allows flexible extended reality (XR) prototyping. The framework includes the planning phase, characteristics, and components needed for the prototypes. The framework's modular design also allows for synchronization in changes across all prototypes. In total, we developed four prototypes, from on-site stadium usage to a virtual reality prototype, where development and evaluation are made possible off-site.

We then investigate the visualization aspect of *ARSpectator*. The main visualization technique we focused on is situated visualization — a method where we present visualizations in spatial relevance to their referents. Based on related frameworks, we developed a conceptual situated visualization framework for on-site sports spectating. Building on that, we implemented and evaluated two situated visualization methods — *Situated Broadcast-styled Visualization* and *Situated Infographics*. Both visualization methods received positive feedback during a user study that we conducted.

Experience from the development of the prototypes showed that technical factors, such as registration, latency, and jitter, impact the user experience. Based on previous work, we investigated three common technical factors — *latency, registration accuracy* and *jitter* to find out the noticeable and disruptive effects they have on user experience. We conducted an experiment in which we highlighted the importance of reducing the effects of these technical factors, as when compounded, there is a considerable disruption to the user experience.

During the development and evaluation of the visualizations, we realized that regardless of how intuitive the visualizations are, an advanced user interface is required for a good experience interacting with the visualizations. Hence, we proposed a context-aware state inference model to analyze the user context. We developed and evaluated a *manual trigger interface* and an *adaptive interface* with potential end-users. Although the concept of a context-aware interface is compelling to participants, our research shows that the interface would need to be well-designed to avoid distractions.

Finally, we proposed a *Stadium of the Future* vision that explains how *ARSpectator* will play a significant role. We also explore the potential of XR technology in providing an interactive experience not only on-site but also for remote spectating. We included ideas that were brainstormed but were not implemented. We then conclude with the future outlook of this research area.

Co-Authorship Form

This form is to accompany the submission of any thesis that contains research reported in co-authored work that has been published, accepted for publication, or submitted for publication. A copy of this form should be at the front (after the thesis abstract) of the thesis submitted for examination and library deposit.

Expectations of theses including publications:

The thesis must be an integrated and coherent body of work. The body of the thesis may include entire publications verbatim, but these need to be reformatted to a consistent chapter style, and commentary may need to be added to link the publications together. If the thesis contains multiple publications, candidates need to ensure that they have: included an introductory chapter; demonstrated their ability to critically engage with the literature; carefully considered their research design and justified choices of methodology and methods where necessary; and synthesised and discussed the findings from the publications. It is the candidate's responsibility to ensure that any published work (or parts thereof) included in the thesis comply with the copyright provisions of the publisher and that any guidelines with regard to self-citation are followed.

Details of publications included in and/or appended to this thesis (please add rows as needed)**.**

Chapter/ Append.	**Paper title**	**Authors**	**Contribution of candidate and co-authors – please detail the nature and extent (%)**	**Journal**	**Status (e.g. under review, forthcoming, published)**
1 & 4	Stats on-site—Sports spectator experience through situated visualizations.	Lo, W. H., Zollmann, S.,& Regenbrecht, H.	Contribution: 85% For this publication, I proposed the design framework, did the implementation, user studies, result analysis, and most of the writing. My co-authors assisted in giving verbal feedback regarding the methodology, discussion of results	*Computer & Graphics, Elsevier (2021)*	Published

			and proof reading of the publication.		
3	From off-site to on-site: A Flexible Framework for XR Prototyping in Sports Spectating.	Lo, W. H., Zollmann, S.,& Regenbrecht, H.	Contribution: 90% I completed the development of the various prototypes and did most of the writing. My co-authors assisted in giving verbal feedback, brainstorming ideas, results discussion and proofreading of the publication.	*36th International Conference on Image and Vision Computing New Zealand (IVCNZ) (2021)*	Published
3	From Lab to Field: Demonstrating Mixed Reality Prototypes for Augmented Sports Experiences.	Lo, W. H., Zollmann, S., Regenbrecht, H.,& Loos, M.	Contribution: 85% I completed the development of the prototype (built on Moritz's work) and did most of the writing. My co-authors assisted in giving verbal feedback, brainstorming ideas, and proofreading of the publication.	*The 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry (2019)*	Published
3	XRSpectator: Immersive, Augmented Sports Spectating.	Lo, W. H., Zollmann, S.,& Regenbrecht, H.	Contribution: 90% I completed the development of the VR prototype and did most of the writing. My co-authors assisted in giving verbal feedback and proofreading of the publication.	*Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology. (2021)*	Published
5	The Impact of Technical Factors on User Experience in Augmented	Lo, W. H., Regenbrecht, H., & Zollmann, S.	Contribution: 85% For this publication, I did the implementation, user	*Computers \& Graphics, Elsevier (2022)*	Under Submission

	Reality Sports Spectating.		studies, result analysis, and most of the writing. My co-authors assisted in giving verbal feedback regarding the methodology, discussion of results and proof reading of the publication.		
5	Technical Factors Affecting Augmented Reality User Experiences in Sports Spectating.	Lo, W. H., Regenbrecht, H., & Zollmann, S.	Contribution: 85% For this publication, I did the development of prototypes and most of the writing. My co-authors assisted in giving verbal feedback and proof reading of the publication.	*Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology. (2021)*	Published
6	A Context-aware Interface for Immersive Sports Spectating.	Lo, W.H., Regenbrecht, H., Ens, B., and Zollmann, S.	Contribution: 85% For this publication, I completed the implementation, user studies, result analysis and most of the writing. My co-authors assisted in the brainstorming process, discussion of results, and proofreading of the publication.	*21st IEEE International Symposium on Mixed and Augmented Reality (2022)*	Under Review
1,2,3	Arspectator: Exploring augmented reality for sport events.	Zollmann, S., Langlotz, T., Loos, M., Lo, W. H., & Baker, L.	Contribution: 20% I contributed to the development of the mobile prototypes explained in the publication. I also created visualizations for all the prototypes and wrote Section 4 - Visualization of the publication. I also assisted with the editing and	*SIGGRAPH Asia 2019 Technical Briefs*	Published

			proofreading of the publication.		
2	Visualization Techniques in Augmented Reality: A Taxonomy, Methods and Patterns.	Zollmann, S., Grasset, R., Langlotz, T., Lo, W. H., Mori, S., & Regenbrecht, H.	Contribution: 10% I assisted in the process of gathering and reviewing related work to be discussed in the publication. Although most of the text is not written by me, I participated in discussions with the main author and did proofreading of the publication.	*IEEE Transactions on Visualization and Computer Graphics (2020)*	Published
3	Augmented Reality for Sports Spectating and Coaching.	Zollmann, S., Langlotz, T., Regenbrecht, H., Button, C., Lo, W. H., & Mills, S.	Contribution: 15% I contributed to the development of the prototypes and assisted with some coding of the user-guided localization approach. I also created the visualizations shown in the prototypes and edited parts of the visualization section.	*Interactive Sports Technologies (pp. 96–111). Routledge. (2022)*	Published

Certification by Primary Supervisor:

The undersigned certifies that the above table correctly reflects the nature and extent of the candidate's contribution to this co-authored work

Name:	Signature:	Date:
Stefanie Zollmann		27/06/2022

Acknowledgements

Firstly, I would like to express my heartfelt gratitude to my incredible supervisors, Stefanie Zollmann and Holger Regenbrecht. They were the backbone support of my PhD, and if given a chance to start over, I would still want them as my supervisors. Every meeting with them inspired motivation and helped me gain confidence in my work while allowing me to learn simultaneously. Besides that, I would also like to thank Chris Button, Stephen-Hall Jones, Barrett Ens, Steven Mills and Tobias Langlotz for contributing to my research.

To my fellow peers in the Visual Computing and HCI group, I thank you for the discussions where everyone was so welcoming and friendly. To my ex-lab mates Lewis, Patrick and Moritz, I appreciated all the help provided. I also want to thank my friends who volunteered to proofread my thesis, namely Lars and Pradeesh, who provided excellent feedback.

I greatly appreciate my friends here in Dunedin, many of whom I have known through Abbey College and the clubs I participated in. Thanks for making my life in Otago vibrant and memorable. To my friends — Pradeesh, Seu Wand, Geraldi, Edward, Alex and Wayne, to name a few, all of you impacted my life during my PhD journey one way or the other.

To my father, thank you for supporting my PhD journey spiritually and financially. I could feel the unconditional love and support despite being so far from home. Thanks to those who kept me sane on social media by chatting with me and sharing a slice of your life back home. Even though I have not met some of you, it kept me in touch with what's going on back there and connected me more to home. Besides that, I sincerely thank my undergrad supervisor, Dr Marina, who saw the value in me and encouraged me to pursue my PhD. I probably would not have considered such an option if it was not for you.

Lastly, this PhD would not have been possible without the support from the MBIE Endeavour Smart Ideas Grant (UOOX1705), which funded my PhD. I would also like to thank our collaborators — Animation Research Ltd., the Highlanders, Otago Rugby Football Union, the Forsyth Barr Stadium and OptaPerform for contributing to this research project.

Contents

List of Figures

Chapter 1

Introduction

Contents

Sports spectating is a popular entertainment for many. However, the on-site sports spectating experience has not seen significant changes since the construction of stadiums. Spectators still go to the stadium to an assigned seat in the stadium to watch a game that happens on the field. There are subtle improvements over the years in on-site sports spectating, such as better quality big screens on-site and smart devices to access information externally when needed. However, the core sports spectating experience has been similar to this point. With the advancement of player tracking (Zhang, Wu, Yang, Wu, Chen & Xu 2020, Blauberger et al. 2021) and mixed reality technologies (Dewangan et al. 2021) in the last two decades, more stakeholders are now interested in improving the sports spectating experience on-site. Spectators themselves have higher expectations of their experience while attending a game in the stadium (Morimatsu 2019). From sports clubs to hospitality venues, improvements in sports spectating experiences would benefit the economy of the city where the sporting event is held (Li et al. 2013). Therefore, our research explores the enhancement of the sports spectating experience on-site with novel technology to achieve a *Stadium of the Future* vision.

To understand the general motivation of this research project, we first present a common sports spectating scenario. Imagine having the opportunity to attend a big game in the stadium for the first time — a final between two popular teams. You arrive at a large stadium crowd and after some effort, you found your seat. Located in the upper corner of the stands, your seat does not have the best view of the action, but you are grateful that you managed to get a ticket anyway. The game begins, and, to your dismay, there was no commentary similar to what you get in sports broadcasts. You could not see very well what was happening on the field, and you could not understand why the referees were blowing their whistles. In the end, you are just there experiencing the atmosphere and have less of an understanding of what is happening compared to when you were watching the qualifiers on television the week before.

Although this might not be the case for every spectator, especially knowledgeable spectators of a sport, it cannot be refuted that little information is provided in an on-site setting compared to watching a sports broadcast. The lack of commentary and game statistics reduces the overall understanding of the game, which is made worse with sub-optimal views. Sports spectators attending live sports events may have the advantage of experiencing the on-site event atmosphere, but it is at the sacrifice of game understanding. Understanding the game is vital to enjoying live sports viewing (Hertzog et al. 2020) and missing out on this could be one factor contributing to a decrease in the number of on-site sports spectators (Koba 2013, Nguyen 2017). To access information similar to that in sports broadcasts, spectators need to look up information from sports statistics websites[1] or mobile applications[2]. However, information from mobile sports applications or websites is usually limited to box score data (a statistical summary of a game) or game meta-data (such as weather or team kit) (Perin et al. 2018). These provide little or no temporal or spatial context for spectators.

[1]https://www.espn.com.au/rugby/

[2]All Blacks Official, OneFootball

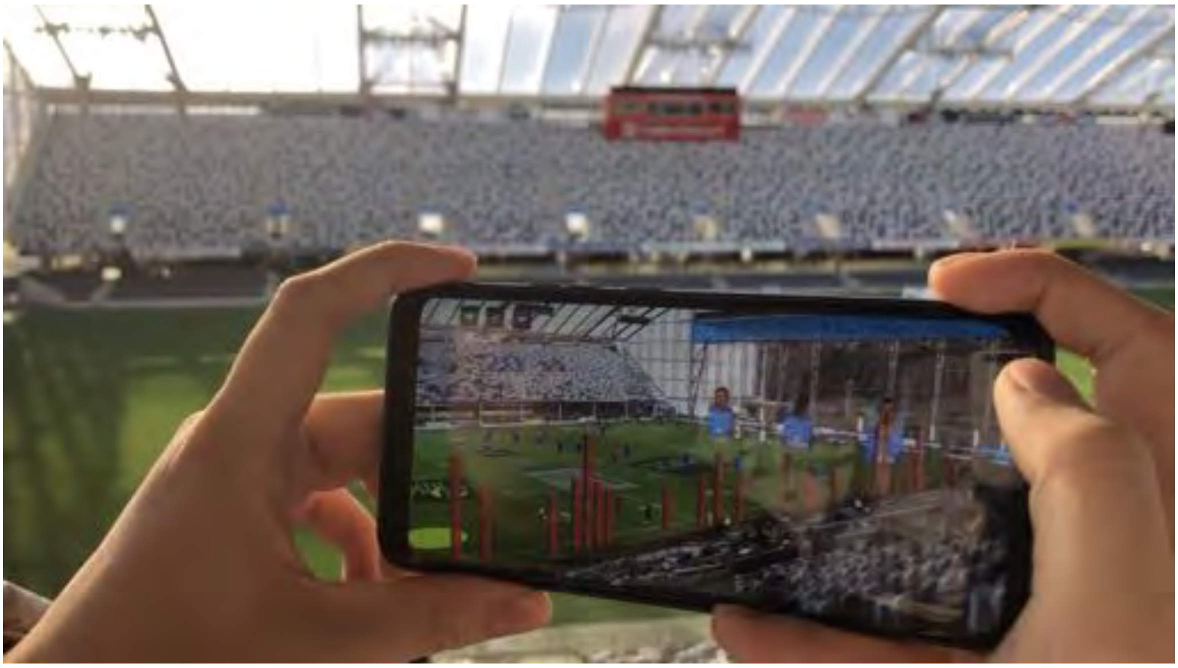


Figure 1.1: Our proposed solution of bridging the gap of sports broadcast and on-site sports spectating - *ARSpectator* prototype on mobile AR.

1.1 Augmented Reality for Sports Spectating

An option to bridge the gap between on-site spectating and remote broadcast experiences is to use **Augmented Reality (AR)** in the stadium (Figure 1.1). AR is a system that aligns the virtual environment coherently with the physical environment (Azuma 1997). AR overlays visualizations onto physical environments to provide more insight. It would provide the best of both spectating medium, providing additional information while retaining the atmosphere on-site. There are already around 3.06 billion Web AR compatible devices as of 2021, with active users reaching 802 million (Boland 2021). However, aspects of using AR for live sports spectating are still under-explored. Soltani & Morice (2020) discussed current applications of AR in sports, ranging from education, spectating, and training using various AR approaches. Other research in this area focus on computer vision techniques to recognize a player from an input image to output primary player profiles (Mahmood et al. 2017). Some researchers use non-real-time broadcast-based systems focusing on post-processing (Cavallaro et al.

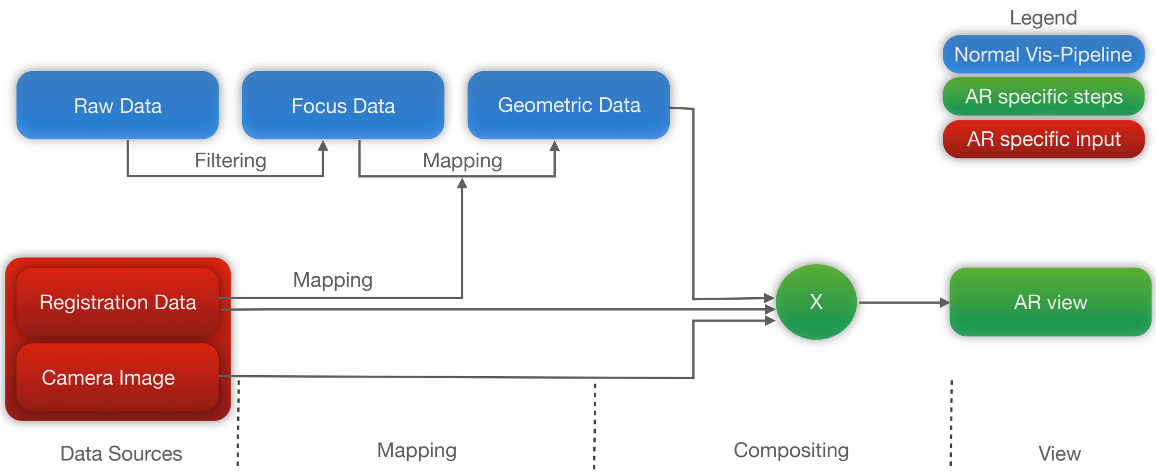


Figure 1.2: AR situated visualization pipeline showing the AR specific steps and input compared to the normal visualization pipeline (Zollmann et al. 2020).

2011). Although there is some interest from commercial companies such as Animation Research Limited[3], Chyron Hego[4] etc., to date, there is no mobile AR system on-site that can provide visual or additional information to spectators.

To facilitate the use of AR in the stadium, we implemented the concept of situated visualizations, which are visualizations placed in the same environment where the referent[5] is in (Willett et al. 2017), such as showing player information on the players themselves.It goes hand in hand with AR and is a core concept of this research. It is the type of visualization we are interested in since it places the visualizations in the field of view of the spectators in the stadium, potentially providing a better spatial understanding. *Visualization pipelines* are a common abstraction that is used to provide key structures in a visualization approach (Moreland 2012). Here, a *visualization pipeline* is shown in Figure 1.2 to illustrate the differences between situated visualizations in AR compared to the generic visualizations. The main differences are the inclusion of camera image and registration data, which leads to a compositing step in which visualizations, paired with the mapping of registration data and camera image,

[3]Animation Research Ltd, URL: https://arl.co.nz/
[4]Chyron Hego, URL: https://chyronhego.com/
[5]The object the visualization is referring to.

are superimposed to get the final AR view.

Essentially, we want to anchor visualizations that spectators might find useful to the environment, eliminating the need for the on-site spectators to search for external information. With that said, there are many questions to be answered, such as where do we place the visualizations, what types of visualization should be shown, and when to present the visualization. This becomes even more complex as sports spectators with different knowledge of the sport would want different types of information. To solve these issues, a context-aware approach is needed. A context-aware application "knows" how to adapt its behavior to a changing environment (Harter et al. 2002), including the users' interest and game state. This would provide suitable visualizations to the correct target audiences in the appropriate time frame. In this thesis, we try to tackle the challenges faced in bringing the experience of using AR in the stadium one step closer to reality.

1.2 Current State of Sports Spectating

In contrast to on-site sports spectating, which has seen less advancement, sports broadcast has seen major technological advancements in recent years. From automatically selecting optimal camera angles for broadcast (Chen et al. 2013), AR experiences in home sports broadcast with the Microsoft Hololens (Stropnik et al. 2018) to the addition of virtual graphical elements in sports broadcast (Cavallaro et al. 2011, Stein et al. 2018), much research and commercial development have been done to provide better experiences for TV or online sports broadcast spectators. Sports broadcast viewers get a better view of the action and gain more insight into the game, all from the comfort of their homes.

Advancements in computer vision also assisted in current-day sports spectating, using multiple high-speed camera setups to achieve accurate ball tracking. Systems such as HawkEye allow tennis balls to be accurately tracked through line extraction, image plane ball tracking, and 3D reconstruction to determine where the ball landed (Owens et al. 2003). Although these systems benefit on-site spectators by providing accurate results, most visualizations are still more beneficial for sports broadcasts. This is because broadcast viewers can see exactly where the ball landed compared to the big screen on-site, where not everyone has a good view of it.

Tsai & Huang (2018) mentioned that a good spectating experience will lead to repeated consumption of sporting events. As we mentioned, there is not a vast change in sports spectating, apart from the easy access to information on the internet which we could now quickly look up if needed through our smartphones. Today, sports arenas offer better facilities and larger screens to display replays and scores. However, the fundamental part of on-site sports spectating remains the same, and few improvements such as having big screens in the stadium have been implemented to understand better the events that occur on the field. We believe it is time for on-site sports spectating to enter a new era of using AR for personalized spatial content. This leads to our research of using situated visualization for sports spectators to enrich their on-site experience.

1.3 AR/XR for Sports Spectating

If AR could bring so many benefits to sports spectators, why is it not widely implemented? There is still much to learn before we can effectively develop situated visualizations for on-site sports spectators. This ranges from localizing the spectators to providing a pleasant, non-disruptive user experience to the spectators in the stadium. This thesis focuses on the latter; we are interested in finding the best methods to provide an enjoyable and usable interface for on-site sports spectators.

This research started with a focus on AR. However, during the research, we realized that we needed different prototypes covering the mixed reality continuum (Milgram & Kishino 1994). Therefore, we expanded the scope of the research to include **Extended Reality (XR)**, an umbrella term for AR, Mixed Reality (MR) and Virtual Reality (VR) technologies (Wohlgenannt et al. 2020). Hence, while in this thesis we mainly refer to AR, many of the contributions are also applicable in XR use cases. Here, we give an overview of the motivation for our **research questions** that we wanted to answer in this thesis before listing them down.

In order to develop the ARSpectator system, we needed to know the components and data sources required so they can be integrated with each other. Hence, this inspired **RQ1**. The formation of **RQ2** initially stemmed from the on-site accessibility of the stadium which the pandemic accelerated the need for more remote experiences, leading to a full suite of remote prototypes. We then realized the need for a framework to develop similar cross-reality prototypes to ensure consistency among all the prototypes.

Since we started our prototyping on mobile, we were looking at different visualization methods. We wanted to take advantage of the spatial understanding provided by the use of AR while also providing spectators with some familiarity. This led to the formulation of **RQ3**, exploring visualization methods that can facilitate spatial understanding. Upon exploring the visualizations, there are then two main components of ARSpectator were not discussed in the thesis - the content sources and user tracking. Therefore, we were interested in researching if these two components would affect user experience and lead to the formulation of **RQ4**.

The final **RQ5** of the thesis revolves around providing a personalized experience, which is a benefit of such AR applications. The question is how we serve the proper content to the target audience seamlessly without taking away too much from the live experience. Hence, in this RQ we developed the Manual Trigger Interface and Context-Aware Adaptive Interface as interfaces to interact with the ARSpectator system.

RQ1. What **components are needed and how do they interact** with each other to support an on-site AR sports spectating use case?

RQ2. How do we **develop XR prototypes for sports spectating** with flexibility and on-site accessibility limitations?

RQ3. What **AR visualization approach** allows for more enjoyable experiences through more informational sports spectating on-site?

RQ4. How do **technical factors in a large-scale environment AR** affect the **user experience** of the sports spectators on-site?

RQ5. How do we provide **contextual information** to spectators without disrupting their sports spectating experience?

Besides the research questions, one of the contributing factors to the lack of AR applications in sports spectating is the accessibility and maturity of AR head-mounted display (HMD) (Billinghurst 2021). AR HMDs are devices that users wear on their heads, allowing them to see the real world with computer-generated graphics augmented onto their field of view. AR HMD could be either optical see-through (OST) where users see the real environment through a transparent screen or video see-through (VST) where the actual environment is displayed on a screen via camera capture. Most AR HMDs are more expensive and less available than VR HMDs. The AR HMDs also have limited battery power, field-of-view (FOV) and portability, making it difficult for the spectator to use them in game. An intermediary solution we are looking at is the use of mobile AR, as many mobile phones now support AR (Boland 2021). However, it does not have all the advantages of AR HMDs, which is the always-on pervasive nature. Mobile AR also introduces arm fatigue after a relatively short usage time.

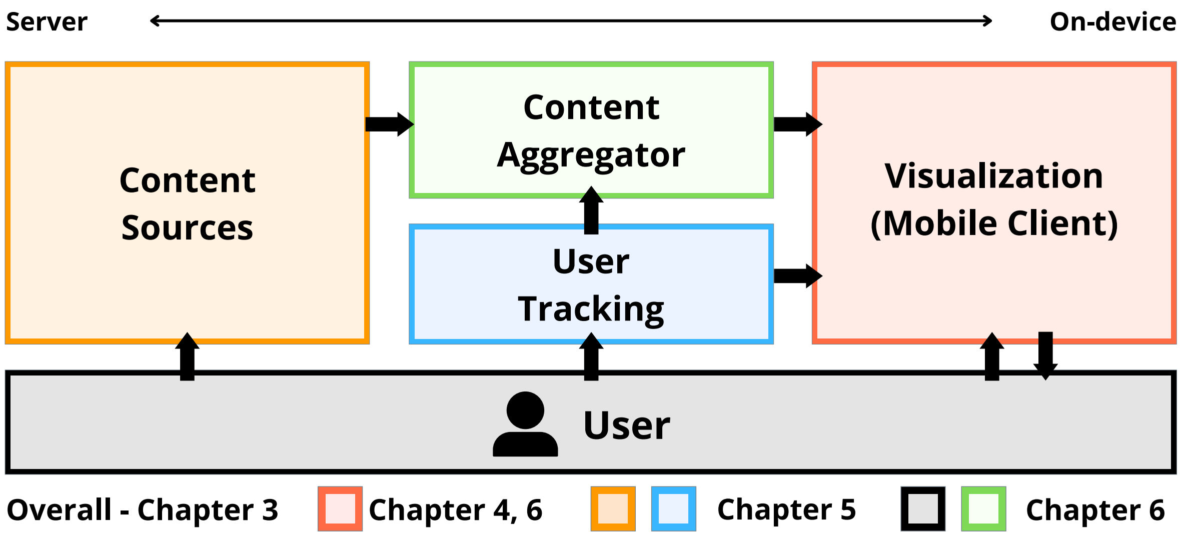


Figure 1.3: A simplified overview of the *ARSpectator* system components and interactions among each other from the server-side implementation to the on-device visualizations. The chapters involved in the components are highlighted.

1.4 Thesis Contributions

This thesis serves as a starting point for the future of on-site sports spectating research in AR. Through our goal of making AR in sports spectating applications possible and enjoyable, we also contribute to sports spectating in XR. The work presented in this thesis mostly surrounds *ARSpectator*'s visualization and user experience aspects, including the development process of different prototypes. Here we illustrate a simple overview of the *ARSpectator* system (Figure 1.3) that shows the components of *ARSpectator*, with server-side content sources up to the visualizations on-device. It also shows which chapters explain the different components and how the user is involved in the overall process. We then summarize the contribution:

1. **An overview** of what components on-site AR sports spectating requires and the interaction between the components. **(Chapter 3, RQ1)**

2. Demonstrating the ***Flexible XR Prototyping Framework*** to overcome obstacles with respect to accessibility to the on-site environment. **(Chapter 3, RQ2)**

3. Presenting **four different XR prototypes developed** during the research with different use cases for each and the reasoning behind why the prototypes were created. **(Chapter 3, RQ2)**

4. Development of the ***3'C's Situated Visualization for on-site Sports Spectating Framework*** that determines where and when to put what type of information in a sports game. **(Chapter 4, RQ3)**

5. **Evaluations of and proposing two situated visualization methods**: *Situated Broadcast-styled Visualization (SBV)* mimicking television broadcast and *Situated Infographics (SI)* where infographics are placed in situ (referent). **(Chapter 4, RQ3)**

6. Investigating the **effects of technical factors on the user experience** — latency, registration accuracy, and jitter in a **large-scale environment AR system**. **(Chapter 5, RQ4)**

7. **Development and evaluation of two interfaces** — A context-aware adaptive interface and a manual trigger interface. (**Chapter 6, RQ5**)

8. Our interpretation of what a ***Stadium of the Future*** would look like and how our research contributes to it. (**Chapter 7**)

1.5 Thesis Outline

An outline of the chapters is provided to give the reader an overview of the flow of this thesis. We also included the publications used in certain chapters of the thesis and the specific contribution that the thesis's author made to each of them.

1.5.1 Chapter Outlines

Here we provide a summary of what the thesis contains, in conjunction with the simplified overview diagram shown earlier (Figure 1.3). Chapter 1 sets the storyline of the thesis. It provides context to the challenges that sport spectators on-site face, which is the lack of contextual information during a live game that leads to decreased game understanding. In this chapter, we also briefly introduce our proposed solution using AR and then outline the contents of the thesis. Chapter 2 presents supporting related work in the overall use case of mixed reality sports spectating to strengthen further the cause of why we conduct this research. A background study provides a baseline understanding of existing research that we can modify and build on.

The following four chapters are the main chapters of the thesis. Chapter 3 provides an introduction and a more detailed overview of the *ARSpectator* system compared to what was shown in Figure 1.3, describing its components and the interaction between them. It also describes the need for a *Flexible XR Prototyping Framework* that we have developed, leading to various prototypes created for this research project out of various necessities. The prototypes are then used in later chapters of the thesis.

Chapters 4, 5, and 6 have their own related work section relevant to the respective chapter. Chapter 4 details our visualization methods, proposing our *3'C's Situated Visualization for on-site Sports Spectating Framework* before introducing two different visualization methods for sports spectating. We conducted two user studies to investigate the user experience of the visualization methods. Chapters 5 and 6 analyze the user experience of such AR sports spectating applications. Chapter 5 looks into the technical factors that could affect the AR user experience, particularly for on-site sports spectating. Chapter 6 then utilizes the framework developed in Chapter 4 and implements it into an adaptive AR interface using context-awareness. We compared the user experience of the adaptive interface with a more traditional manual interface in a user study.

Finally, Chapter 7 summarizes the findings of all previous chapters. We introduce our *Stadium of the Future* concept and describe how our research would fit with this vision. We discuss the potential of *ARSpectator* and what future work could be implemented to bridge the growing gap between sports spectating on-site and sports broadcast.

1.5.2 Declaration and Collaboration

This thesis contains parts extracted from the publications I have published. In this section, I outline the thesis structure and declare the publications included in each chapter. Materials used that are not cited are publications in which I am the first author. I cited all materials that were obtained from publications that I co-authored and listed them below. Although I am the author of this thesis, I use the term “we" in this thesis to be inclusive of the whole research team.

The initial lab-based AR prototype was developed by one of the lab interns, Moritz Loos. The current lab AR prototype was developed by me but was built upon Moritz's work. In addition to that, the 3D Computer-aided Design (CAD) stadium model, including the textures used in this research, were provided by our collaborators Animation Research Limited (ARL). The other visualizations were made in Unity Game Engine[6] with some visualizations designed in Canva[7] before being imported into Unity.

This chapter **Chapter 1** consists of content published as:

[6]https://unity.com/

[7]https://www.canva.com/

1. Lo, W. H., Zollmann, S., & Regenbrecht, H. (2021). Stats on-site—Sports spectator experience through situated visualizations. Computers & Graphics, Elsevier.

For this publication, I proposed the design framework, did the implementation, user studies, result analysis, and most of the writing. My co-authors assisted in giving verbal feedback regarding the methodology, discussion of results and proof reading of the publication.

Chapter 3 consists of content published as conference papers, demonstration abstracts, and posters:

1. Lo, W. H., Zollmann, S., & Regenbrecht, H. (2021). From off-site to on-site: A Flexible Framework for XR Prototyping in Sports Spectating. 2021 36th International Conference on Image and Vision Computing New Zealand (IVCNZ).
2. Lo, W. H., Zollmann, S., Regenbrecht, H., & Loos, M. (2019). From Lab to Field: Demonstrating Mixed Reality Prototypes for Augmented Sports Experiences. The 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry.
3. Lo, W. H., Zollmann, S., & Regenbrecht, H. (2021). XRSpectator: Immersive, Augmented Sports Spectating. Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology.

For these publications, I completed the development of the various prototypes and did most of the writing. My co-authors assisted in giving verbal feedback on the developed prototypes, brainstorming ideas, and proofreading of the publication.

Chapter 4 is edited according to the publication:

1. Lo, W. H., Zollmann, S., & Regenbrecht, H. (2021). Stats on-site—Sports spectator experience through situated visualizations. Computers & Graphics, Elsevier.

This is the same publication that has been mentioned in the declaration for Chapter 1, hence the contribution is the same.

Chapter 5 is based on the publications in Computer & Graphics Journal (under submission) and poster at VRST 2021:

1. Lo, W.H., Regenbrecht, H. and Zollmann, S., The Impact of Technical Factors on User Experience in Augmented Reality Sports Spectating. (Computer & Graphics, Elsevier)

2. Lo, W. H., Regenbrecht, H., & Zollmann, S. (2021). Technical Factors Affecting Augmented Reality User Experiences in Sports Spectating. Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology.

For these publications, I completed the implementation, user studies, result analysis and most of the writing. My co-authors assisted in giving verbal feedback on the methodology, discussion of results, and proofreading of the publication.

Chapter 6 is submitted for publication to ISMAR 2022 (under review):

1. Lo, W.H., Regenbrecht, H., Ens, B., and Zollmann, S., A Context-aware Interface for Immersive Sports Spectating. (Submitted to ISMAR 2022)

For this publication, I completed the implementation, user studies, result analysis and most of the writing. My co-authors assisted in the brainstorming process, discussion of results, and proofreading of the publication.

Finally, here is the list of co-authored publications that are cited throughout this thesis. The contributions I have made to the publications are also detailed.

1. Zollmann, S., Langlotz, T., Loos, M., Lo, W. H., & Baker, L. (2019). Arspectator: Exploring augmented reality for sport events. In SIGGRAPH Asia 2019 Technical Briefs (pp. 75–78).

2. Zollmann, S., Grasset, R., Langlotz, T., Lo, W. H., Mori, S., & Regenbrecht, H. (2020). Visualization Techniques in Augmented Reality: A Taxonomy, Methods and Patterns. IEEE Transactions on Visualization and Computer Graphics.

3. Zollmann, S., Langlotz, T., Regenbrecht, H., Button, C., Lo, W. H., & Mills, S. (2022). Augmented Reality for Sports Spectating and Coaching. In Interactive Sports Technologies (pp. 96–111). Routledge.

For co-authored publication 1, I contributed to the development of the mobile prototypes explained in the publication. I also created visualizations for all the prototypes and wrote Section 4 — Visualization of the publication. I also assisted with the editing and proofreading of the publication.

For co-authored publication 2, I assisted in the process of gathering and reviewing related work to be discussed in the publication. Although most of the text is not written by me, I participated in discussions with the main author and did proofreading of the publication.

For co-authored publication 3, I contributed to the development of the prototypes and assisted with some coding of the user-guided localization approach. I also created the visualizations shown in the prototypes and edited parts of the visualization section.

Chapter 2

Background and Related Work

Contents

There has been increasing interest in using MR technologies to enhance user experience and productivity, probably due to the availability of VR and AR HMDs at consumer electronics price points (Rokhsaritalemi et al. 2020). Mobile AR has also helped drive AR to the masses, as evident in the one billion-plus downloads of the mobile AR game Pokemon GO had (NintendoSoup 2019). This form of AR opens up research opportunities in various fields, including our use case of live sports spectating.

Although our research began with the intention of using AR as our primary medium, we slowly delved into VR as a means of assistance for the times when we had limited access to on-site facilities. We have already briefly gone through some of the core concepts of AR and VR in the previous chapter. In this chapter, we first provide examples of some AR/VR use cases. Then, we will dive deeper into AR/VR visualization methods and design frameworks before touching on some user experiences.

2.1 Extended Reality

Extended reality (XR) is an umbrella term for AR, MR and VR technologies (Wohlgenannt et al. 2020). Some researchers (Palmas & Klinker 2020, Mann et al. 2018) summarized that the 'X' in XR could mean differently, from x as a variable on the reality - virtuality continuum, including future technologies, to 'X' as a cross-over between different technologies. For the scope of this research, we focus more on the use of AR and VR. Therefore, we refer to our prototypes as XR prototypes, as they span between different types of AR and VR.

2.1.1 Augmented Reality

The term "Augmented Reality" can be traced back to 1968, when Sutherland (1968) developed the first AR system. It has been defined multiple times by various researchers in different fields. One of the most cited definitions is by Azuma (1997), Azuma et al. (2001) where AR is classified as a system that combines and aligns real and virtual environments, contains real-time interactions, and 3D registrations of objects. These characteristics do not impose any technology limitation, be it an optical see-through (OST) or video see-through (VST) HMD, optical projection, or mobile. OST devices allow users to view the actual environment directly with little or no perspective distortion (Cutolo et al. 2020), while VST uses the camera on the device to view the environment. Other definitions of AR could be briefly summarized as *"An approach that bridges the gap of the real and virtual environment by using and adding digital contextual information to one's natural environment, thus augmenting the users' experience through invisible computer interfaces."* (Klopfer & Squire 2008, Billinghurst et al. 2015, Sanna & Manuri 2016).

We could use AR applications in various domains. Spohrer (1999) introduced the concept of WorldBoard roughly 20 years ago—a vision about "*putting localized information objects in physical space on a planetary scale grid*". It describes a world filled with geo-coded visualizations integrated into our daily lives. This concept is closely related to AR, as Spohrer (1999) suggested using AR as an interface for WorldBoard. However, large-scale implementation of the concept was difficult to achieve at that time due to technical limitations, for example, positioning and tracking (Spohrer 1999, Azuma 1993).

Hence, we could provide some popular AR use cases by borrowing some of the applications mentioned in the WorldBoard. One of the notable fields in which AR could thrive is education. Research on AR in education (Bower et al. 2014) includes location-based AR (Billinghurst & Duenser 2012), which brings benefits such as long-term memory retention, better understanding and motivation alongside the ease of collaboration (Radu 2012). Education could possibly be the field to utilize AR's full potential soon. The study of using AR in medical field training is also gaining traction. While the technology has not evolved to the mature point where it can be implemented for actual medical use, it serves as a good potential training platform, as the situated learning experience might be transferable to an actual workplace (Barsom et al. 2016, Kamphuis et al. 2014).

Another target application for AR would be industrial, infrastructure, and public safety. For industrial use cases, it is more often in the form of maintenance and repair where instructions could be provided in spatial relation (Kohn & Harborth 2018, Bellalouna 2020). In recent years, public infrastructure application scenarios have been explored as mobile AR applications, e.g. to prevent damage to underground utilities while conducting maintenance or new development (Schall et al. 2010, Talmaki et al. 2010). Another example is AR-based navigation systems for car drivers and pedestrians (Jang & Hudson-Smith 2012, Narzt et al. 2006, Grubert et al. 2016), which would provide better safety as users of the road do not need to distract their attention elsewhere.

The AR use cases not only cover what we said above, which is more towards the serious and critical matters. It also includes tourism, recreation and sports. Many AR applications are being implemented in tourism to promote interactive learning, especially when it involves cultural heritage with a focus on user-centered design (Cranmer et al. 2020, Jingen Liang & Elliot 2021). AR has also been widely used in gaming, especially mobile games (Liberati 2018). The abovementioned Pokemon GO allows players to place virtual characters into the players' environment using mobile AR (NintendoSoup 2019). Our research focuses on AR in sports from a spectator's perspective. A majority of the work done so far is in the field of training, performance analysis and broadcasting (Cavallaro et al. 2011, da Silva et al. 2021). These are either for broadcast spectators or are more targeted towards the players or coaches themselves for the use of performance evaluation, rather than standard consumption for sports spectators. Hence, this thesis fills the gap of contributing towards the on-site sports spectators.

Mobile AR

Mobile AR is a suitable medium for *ARSpectator* until AR HMD becomes more affordable to the masses. The reason is the abovementioned availability of AR-compatible mobile devices (Boland 2018), where almost everyone with a modern mobile phone can install mobile AR applications. The mobile AR that we refer to here is the use of AR via a mobile wireless device, mainly through a smartphone or tablet (Kourouthanassis et al. 2015, Chatzopoulos et al. 2017). One could argue that an AR HMD like the Microsoft Hololens is also a mobile wireless device, but for this thesis, we will not include that as mobile AR and classify that as a typical AR application. Mobile AR is simply a form of VST AR where users see the environment through the lenses of the mobile device. Visualizations are then overlaid onto the screen in spatial relation to the environment, usually through computer vision methods (Kim et al. 2014, Amin & Govilkar 2015) but more recently with the aid of LiDARs for shorter range AR (Tavani et al. 2022).

Indirect AR

Indirect AR uses a pre-recorded 360° image/video of an environment to simulate an AR experience. Wither et al. (2011) describes indirect AR as a method of minimizing visual disturbance, often caused by instability or imperfection of localization and tracking. We are interested in indirect AR as it allows us to develop and experience the AR experience off-site. For some parts of our research, we focused on a mobile version of indirect AR due to the lack of access to the stadium, coupled with the restrictions of the COVID-19 pandemic, to conduct off-site testing and user studies. However, indirect AR does come with the challenges of spatial and temporal inconsistencies, as mentioned by Skinner et al. (2018) which will lead to a slight decrease in immersion.

2.1.2 Virtual Reality

VR is the simulation of an entirely artificial virtual environment, providing full virtualization (Buhl & Winter 2009) to the viewer. This multi-sensory experience provides a sense of immersion to the user as if they are in the virtual environment (Bowman & McMahan 2007). There are also many other benefits of VR, such as the opportunity to manipulate scenes in a way that is impossible in the real world, from weather control to a third-person perspective of oneself (Petri et al. 2018, Evin et al. 2020). Similarly to AR, VR also rose in popularity due to VR HMD becoming more affordable and available, although it has been around for a long time (Wohlgenannt et al. 2020).

VR is often used for simulation training, particularly in the medical field, as summarized in the survey paper by Li et al. (2017). It is also widely used in the industry (Román-Ibáñez et al. 2018) and education (Radianti et al. 2020), where institutions could provide simulated hands-on experience without purchasing expensive equipment. An example of this would be simulation of surgeries, where guidance could be provided while the practitioner could physically perform the action required (Lungu et al. 2021).

Just like indirect AR, VR is used in this research project as it is a handy tool to simulate AR without the interference of tracking issues. Rather than mobile indirect AR just having the stadium environment on the mobile phone screen, we could now immerse the user in a complete stadium environment, even when we are not at the stadium. Such AR in VR methods has also been applied in medical laparoscopy (Hettig et al. 2018). However, we could also apply VR to other sports spectating use cases, such as remote spectating, which could be particularly useful during the pandemic, as shown by Singh et al. (2020).

2.2 Situated Data Representations

Situated Data Representations is the representation of data in a given space. Visualizations in AR are different compared to standard *Information Visualization* and

Scientific Visualization. Information visualization uses computer-supported interactive methods to amplify cognition (Card 1999). On the other hand, scientific visualizations are similar but are based on simulation outcomes or measurement data from which researchers could extract knowledge from it (Brodlie et al. 2012). The physical world is not relevant to the visualizations in both of them. Visualizations in AR take the main form of situated data representation (Willett et al. 2017), an umbrella term for a visualization method where visualizations are close to the physical referent of the data. This includes a few different visualization methods used in AR, and in the following subsections, we will explore some common situated data representation methods. Although situated data representation is different to traditional visualization methods, it comes with similar challenges to the traditional visualization method, such as cluttering of information (Ellis & Dix 2007), visual scalability (Liu et al. 2014), and choosing the proper visualization for that particular data.

2.2.1 Requirements for Situated Data Representations

The requirement for situated data representations is to have close spatial proximity between the data presented and the physical referent in the actual environment (Willett et al. 2017). It echos White & Feiner (2009*b*) definition of situated visualization where visualizations have relevance with the environment in which they are placed.

2.2.2 Types of Situated Data Representations

Situated Visualization

The majority of the visualizations in this thesis will be in some form of situated visualizations. White & Feiner (2009*b*) introduced situated visualization as a concept of visualizations in a spatial context, a combination of visualizations and their relationship with their environment. An example of this would be to display the stock count of items in the warehouse itself to get an easy grasp of what items are running low. Later, Willett et al. (2017) provided a similar definition of visualizations simply being in the same space as the referent, though primarily just visually on-screen. In contrast, visualizations that have data closer to the referent are known as embedded visualization. These are the two most popular definitions used in the past literature according to a survey by Bressa et al. (2021). Both describe visualization techniques displaying visualizations relevant to their location or physical context. For this thesis, we take situated visualizations as the umbrella term for our visualizations, which includes the subset of embedded visualization.

With the above definition, some might think that all AR visualization would be some form of situated visualization, which is incorrect. Suppose a person is using a visualization toolkit (Sicat et al. 2019) to visualize augmented data on a tabletop and is not related to the environment in which it is located. In that case, it is not situated visualization. However, if an AR system augments instructions overlaid onto the screen space, it will still be considered a situated visualization even though there is no anchoring to the environment.

Situated Analytics

ElSayed et al. (2015) introduced situated analytics as a combination of visual analytics (Keim et al. 2008) and AR to embed visual representations into the physical environment (Thomas et al. 2018). This object-centered approach is interactive and data-oriented. An example of this includes showing nutritional facts augmented on food items and probably rank the items in terms of how healthy it is. The main goal is to support users in navigating a multi-dimensional database, including ranking, filtering, and locating physical objects based on queries. Situated analytics also deals with similar challenges of situated visualizations, such as clutter management and intuitive visualization methods (ElSayed et al. 2016*b*). In the example given by ElSayed et al. (2016*b*), users on mobile devices explore the nutritional value of different cereals based on what they select. However, this technique focuses exclusively on selected objects of interest. There are no forms of interconnected entities in this situation other than which cereal is the best in a particular category that the users choose.

Immersive Analytics

Immersive analytics (Chandler et al. 2015) combines user interactions and multi-sensory systems to place users in an engaging and immersive environment for data analysis. Although immersive analytics and situated analytics have appeared around the same time, there are differences. Unlike situated analytics, immersive analytics does not require data representation to be relevant to its environment. Immersive analytics opens up a new dimension of exploring and manipulating data in 3D space while allowing collaboration with others in the same virtual environment. We could implement immersive analytics everywhere, such as entirely virtual in a VR environment, in a sizeable CAVE Automatic Virtual Environment (CAVE)-like setting or even in AR (Cordeil et al. 2019, 2016, Herr et al. 2017).

2.3 Existing Sports Visualizations

Sports visualization for broadcasting and web-based applications has advanced considerably in recent years. Perin et al. (2018) compiled a list and reviewed recent work on the visualization of sports data based on the classification of sports data from box score data, tracking data, and metadata. Box score data is the structured summary of the game while tracking data are any dynamic data that involves positioning. Metadata is data that describes other data, summarizing basic information about a particular subject. When applied to an example in our rugby use case involving a player scoring a penalty, the box score would be the score itself and previous penalty stats. The tracking data could be the position the kick was made, and the metadata could be the player's expected performance (prediction). However, most of the current work done is of analytics and is not in real-time, let alone on-site AR. These visualization types are more suitable for a tabletop AR scenario or a virtual experience in which graphs and charts could be drawn and explored in 3D space, such as by using the DXR toolkit (Sicat et al. 2019) or the Immersive Analytics Toolkit (Cordeil et al. 2019). There are also more general data classifications, such as the primary classification by nominal, ordinal, and quantitative (Card & Mackinlay 1997) and the multi-dimensional classification by Shneiderman (1996).

We got inspiration from existing literature and commercial sports broadcast visualizations to systematically develop AR-specific sports visualization methods. This section will introduce some of the commercially done sports visualizations, not limited to rugby. The existing visualizations are then questioned whether they would be suitable in the context of sports AR and what changes could be adapted to suit the characteristics of AR. We then examine the stadium model to find all the possible canvases where visualizations could occur, which allows us to analyze where a specific visualization method could be applied.

2.3.1 AR Sports Visualization

There is a growing interest in AR situated visualizations for sports (Zollmann et al. 2019, Lin et al. 2020). SportsXR (Lin et al. 2020) identified the potential for coaches, fans, and even players in terms of training but also mentioned some technical challenges such as data collection and visualization design. Companies such as Immersiv.io[1] and Nexus Studios[2] showcased on-site AR applications for soccer and basketball. However, apart from demo videos and showcases, there is not much more information available publicly yet. Our observation is that the research area of on-site AR in sports spectating is still emerging. There is only a little previous work, with most of them done in terms of coaching and training rather than the spectators' experience.

Previous AR-based sports application research focuses mainly on player identification through image processing (Mahmood et al. 2017, Bielli & Harris 2015), off-site AR-based broadcasting (Cavallaro et al. 2011) and off-site AR used concurrently with live broadcasting (Stropnik et al. 2018). There is also research on using AR for gamification and social reaction sharing in sports (Thompson & Potter 2017). However, what is missing overall is a conceptual framework for situated sports visualization, as most researchers focus on a specific technical implementation. Our work provides a basis for all situated sports visualization, from where to place content to how to display it.

[1] https://www.immersiv.io/

[2] https://nexusstudios.com/work/samsung-ar/

2.3.2 Analytical Sports Visualization Tools

Several studies developed tools for usage in a game that implements situated visualization but is not an AR application. The application relevant to our research is MatchPad (Legg et al. 2012), an interactive glyph-based visualization tool for real-time sports performance analysis. Although the information is presented on-screen in a traditional manner, it presents real-time sports data in the form of glyphs (Chung et al. 2015). This allows users to quickly understand the context of what is going on, including a few optional details, all in a scale-adaptive timeline layout. SoccerStories (Perin et al. 2013) provides a visual analysis of a soccer game, allowing users to see much information at a glance while having temporal zoom towards specific game phases for time-specific details. Such applications have also been made in other sports, such as baseball (Lage et al. 2016). There are also researches combining sports broadcast videos with other data sources to provide analysis for a team sport (Stein et al. 2018).

2.3.3 Sports Broadcasting

Animation Research Limited (ARL)[3] is a production house that turns digital data into visualizations. A division of ARL focuses on sports broadcast visualizations called Virtual Eye. Virtual Eye specializes in creating real-time sports graphics on television, recreating virtual scenes to provide home viewers with a different perspective of the game they are watching. Although usually done for television broadcasts and panoramic videos, some of their visualization methods could also be implemented in AR. Examples of it include a floating billboard to show statistics (Figure 2.1(a)) or a picture-in-picture replay (Figure 2.1(b)) where they augmented replays into the environment. They also used multiple labeling systems and ball path trajectories, providing viewers with a better game summary.

Similarly to ARL, Chyron Hego[4] also does computer graphics for sports broadcasts. Making use of chroma key compositing technology (Shimoda et al. 1989) in their Virtual First product, they create spectacular virtual content placement, from the product ad

[3]Animation Research Ltd, URL: https://arl.co.nz/

[4]Chyron Hego, URL: https://chyronhego.com/

Figure 2.1: From top left (a) Floating Billboard, (b) Picture-in-Picture Replay, (c) Virtual Placement, (d) Highlighting of Players with labels, (e) Virtual Player Profile, (f) Ball Path Trajectory, (g) Smart Vote System, (h) On-field annotations

on the field (Figure 2.1(c)) to the annotations made by an operator in real-time. The method of highlighting players on the field is suitable for AR as it easily enables users to identify a specific player of interest. However, chroma-keying might be slightly more of a challenge for an AR app as the broadcast camera is calibrated in advance to provide close to perfect virtual content placement. Virtual First also allows the labeling of players and annotations in the field, which was also of interest (Figure 2.1(d)).

Hawk-Eye Innovations[5] specialize more in officiating games. With an array of up to 30 automated cameras, they could quickly determine if a ball crossed the goal line or a foul occurred. This setup provides an accurate and quick decision from the operator independently of the host broadcaster. Hawk-Eye also creates some visualizations, such as ball trajectories (Figure 2.1(f)) and virtual player profiles (Figure 2.1(e)) augmented to look as if they were there in the real environment. This is highly useful in an AR application as a distinct difference in graphics quality would disrupt the realism of the user AR experience. Hawk-Eye also deployed a voting system (Figure 2.1(g)) in which spectators in a live match could engage by holding special objects of different colors to be detected by the cameras. The results are then shown on the screen in the stadium and augmented on the screen for home viewers.

[5]Hawk Eye, URL: https://www.hawkeyeinnovations.com/

Lastly, we have Vizrt[6], a company that also creates computer graphics for television programs such as news and sports broadcast. Using data-driven AR graphics, they create virtual scenes that are used more toward analytics to dissect events that occurred in the game. With the use of multiple cameras, they could replay a clip and transition to a different point of view, creating the virtual effect of being in the stadium. In addition to that, players in videos could also be cropped and moved around, significantly improving the storytelling experience. Vizrt also creates many aesthetically beautiful visualizations such as player introductions, game field annotations, and many more (Figure 2.1(h)).

2.4 Design Frameworks and XR Prototyping

2.4.1 Taxonomy and Frameworks of Existing Situated Visualization Concepts

A visualization framework is an essential base for generating similar visualizations reliably and effectively in the future. A few frameworks are relevant to our goal of supporting on-site spectators in a stadium environment. One is the AR-CANVAS framework for embedded visualization (Bach, Sicat, Pfister & Quigley 2017). This framework discusses key terms, parameters, and challenges that should be considered when designing an embedded visualization. In a co-authored paper Zollmann et al. (2020), we developed a visualization taxonomy and framework for AR in general, which shows the components and data flow in the visualization process. White & Feiner (2009*b*) also developed an AR visualization framework in which the author characterizes various AR applications according to context, relevance, display type, presentation, and interaction.

We extended White & Feiner (2009*b*) classification with our *ARSpectator* system (Figure 2.2). The orange parts represent the current implementation and the blue parts represent planned future implementation. Those that are white are not fitting in this *ARSpectator* project, but is still relevant to the study of AR. Our implementation

[6] Vizrt, URL: https://www.vizrt.com/

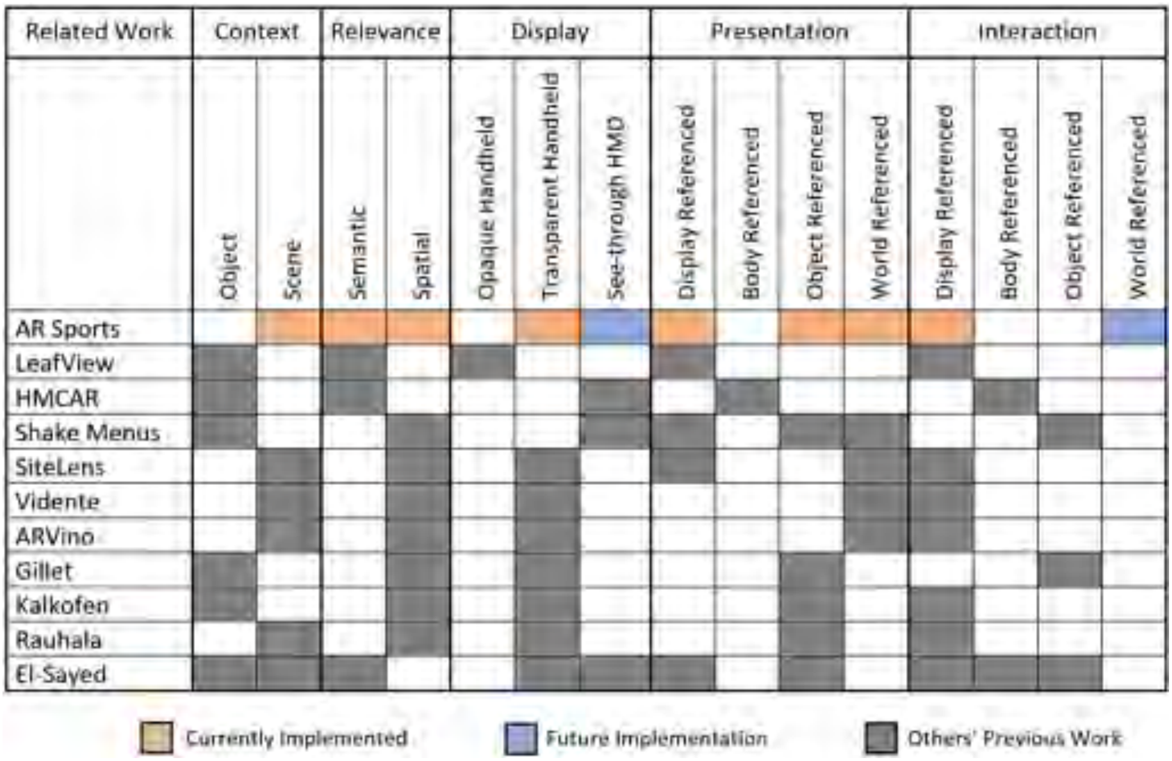


Figure 2.2: Characterization of our application using White & Feiner (2009*b*)'s classification framework. The table is obtained from the White & Feiner (2009*b*) technical report with the addition of our proposed work and the work of ElSayed et al. (2016*a*)

involves a stadium environment to integrate content, but we will discuss this in more details in Chapter 3. According to White & Feiner (2009*b*) classification, our implementation includes the scene context and uses both semantic (the players on the field) and spatial relationships (related to where events happen on the field). Our research also aims to provide a world-referenced interaction, using the environment to interact with the data, which is a first according to White & Feiner (2009*b*). Additionally, we added an entry on Situated Analytics (ElSayed et al. 2016*a*) as it also involves situated visualization and explores multiple objects in a scene. However, Situated Analytics focus on a smaller scale and requires more complex user input, rendering it unsuitable for our scenario.

Infographics

Infographics are the visual representation of data that aims to reduce data complexity. In addition to the attractive form and effective representation (Otten et al. 2015), infographics provide visual learning to aid the short attention span people have with the overload of information (Smiciklas 2012). This, together with the visual nature of infographics, has the potential to support data visualization in AR — complex data will be shown in a comprehensible visual format, and spatial relationships are provided by the AR concept itself, hence implementing a visualization concept of the World-Board. At present, there are three main categories of infographics: *static infographics*, *animated infographics*, and *interactive infographics*.

Static infographics traditionally come in the form of infographics seen on posters, webpages and articles. Siricharoen (2013) listed the main types of infographics based on the data presented, mainly statistical, timeline, process, and location or geography-based infographics. Other studies emphasised the importance of highlighted color, flow and complexity in infographics (Harrison et al. 2015, NuhoKibar & Akkoyunlu 2017). Some studies investigate how infographics could reduce the language barrier and persuade the readers better than traditional text and images (Siricharoen 2015, Lazard & Atkinson 2015).

Animated infographics are created for screens in a video format where there is motion and animation for the elements and data in the infographic (Hassan 2016). Mayer et al. (2005) investigated the potential of using animated infographics compared to static infographics to illustrate processes that are not visible in a real environment, such as the movement of particles to form clouds. Afify (2018) explained that animated infographics provide more information to the user, but possibly at the cost of sacrificing flexibility in consuming the data as users cannot choose what they want to see. Therefore, *interactive infographics* were developed, infographics that allow users to control the graphics in a perceptible way (Weber 2017). For example, thredUP, a sustainable fashion brand, used a quiz-like interactive infographic to explain the carbon footprint from fashion[7]. Despite the widespread use of infographics, users need to

[7] https://www.thredup.com/fashionfootprint/

spatially link the information presented in the infographics to the topics of interest. Providing such a spatial link of information is a significant component of our research.

Immersion

Immersion is the objective level of sensory fidelity that a VR system provides (Slater 2003). It is one of the advantages that mixed reality information visualization has over traditional information visualization (Zheng et al. 2017, Scholz & Smith 2016, Bach, Sicat, Beyer, Cordeil & Pfister 2017, Sicat et al. 2019). It would be desirable to have visualizations implemented in a way that does not distract us from the reality in front of us. While VR is immersive due to its full view coverage, AR applications adapt immersion by having unobtrusive visualizations that blend well with the environment, such as ghosted views, geometric guidance, highlighting objects, and aligned label management (Kalkofen et al. 2013, Keil et al. 2018, Tatzgern et al. 2014). These subtle visualizations tend to blend into the environment, reducing the jarring difference in visuals of the visualizations and environment. This requires, for example, occlusion management, photorealistic rendering and high-quality 3D reconstruction (Rohmer et al. 2017, Osti et al. 2019, Agusanto et al. 2003). For example, recent ARCore[8] developments by Google illustrate the possibilities of occlusion management of AR content in a real environment via their Depth API (Izadi 2019).

Visualization Components

Several visualization components developed are relevant to discuss within the visualization of information in AR. Munzner (2014) provided an overview of what visual marks and channels could do to create a different perspective in a visualization. We mentioned glyphs earlier, which help visualize multivariate data in a compact form (Ward 2002, Chung et al. 2015), making them suitable for attaching to small spaces. Tatzgern et al. (2016) used glyphs to provide an overview of the books and developed an adaptive algorithm to hide or display visualizations to prevent cluttering. Horus Eye (ElSayed et al. 2016*a*) uses colors to create contrast among objects that are compared

[8] ARCore, https://developers.google.com/ar

using opacity as presented by Livingston et al. (2003).

2.4.2 Prototyping in XR

Wensveen & Matthews (2014) classified four prototype roles; an experimental component, a means of inquiry, a research archetype, and prototyping as a vehicle for inquiry. This classification did not include prototypes developed as a tool to facilitate the development and research process itself. Meanwhile, Koskinen & Frens (2017) argued that research prototypes and design or industrial prototypes are different since research prototypes are meant to test theoretical literature rather than looking more *"product-like"*. We think our prototypes are somewhere in the middle of both concepts, a *"product-like"* prototype yet used as an experimental component. One could also call this a minimally viable research product.

There are low-fidelity and high-fidelity prototypes (Lim et al. 2006). Typical examples of low-fidelity prototypes are sketches and wireframes that require users' imagination to fill in the fidelity gaps. Sometimes this can be augmented by the Wizard of Oz technique (Bernsen et al. 1994) where the facilitator manipulates the system while a subject is interacting with it. High-fidelity prototypes are closer to fully functional products where users can interact with them. Mixed-fidelity prototypes are also closer to high-fidelity prototypes but still have some manual elements in them (De Sá & Churchill 2012). Among all these classes of prototypes, De Sá & Churchill (2012) found that although the low-fidelity prototype is the easiest to produce, it has the lowest score when it comes to understanding the concept of the prototype.

Indirect AR (Wither et al. 2011) is one of the approaches we used for our prototypes that provided many benefits. While playing a video on-site (De Sá & Churchill 2012) might be similar to an indirect AR approach, it does not provide the freedom for users to look around. Apart from just simulating an AR experience, we could use indirect AR in situations where the crowd might be a factor that affects user experience. An example is this museum use case where a combination of traditional AR and indirect AR allows users to see the environment without the occlusion of other visitors (Gimeno et al. 2017).

2.5 Evaluating the XR User Experience

We are interested in providing a better user experience during on-site sports spectating by providing better insights through situated visualization (Tatzgern 2015) to spectators. Prior work in AR focuses mainly on the influence of human factors (Livingston 2013) or technical aspects such as optimizing the performance of a tracking algorithm or display properties. Human factors involve engineering a certain product in a way that is suitable for use by the intended target audience, for example, an excellent safe design or an exemplary user interface that leads to effective use (Chapanis 1991). In AR research involving user studies with human participants, performance measures such as goal achievement, task completion rates, ease of use, and knowledge gain (Rengger 1991, Livingston & Ai 2008) are used. In contrast, technical papers propose novel algorithms and techniques, for instance, tracking, registration and novel display types. These technical papers often focus on technical performance measurements, such as tracking rates, accuracy, re-projection errors, or pose errors (Gao et al. 2017), without investigating the impact of these performance measurements on user experience. Until now, the investigations on human factors and technical factors have often been separated.

A survey of evaluation techniques from the early 1990s to the late 2000s shows that there are not many user evaluations in AR research in this time frame (around 10%) (Dünser et al. 2008). Only since the early 2000s, there has been a shift in focus toward research in AR user experience (Poppe et al. 2007), and there have been calls to consider some of the technical factors such as tracking into the design phase (Mulloni et al. 2012). However, to the best of our knowledge, little to no research investigates how these technical factors affect the user experience.

Chapter 3

The ARSpectator System

Contents

The lack of an implemented solution for on-site sports spectating mentioned in the introduction prompted us to develop a prototype system that supports the use of AR in an on-site sports spectating scenario. We call our developed system *ARSpectator*, a combination of using AR as a spectator in the stadium. In this chapter, we will go through the overall concept of *ARSpectator* (**RQ1**), our development process in which we introduce a flexible XR prototype development framework (**RQ2**), and details on our four different prototypes (**RQ2**), which we will use throughout the thesis.

The main goal of *ARSpectator* is to have an improved experience during a live game in the stadium. Although some of the technology is not yet present to implement our ideal use case, we try to describe the vision of *ARSpectator*. The ideal use case for *ARSpectator* is where spectators can get useful visualizations when they want or need to without much effort. It does not distract them from the game and provides additional engagement with the crowd. Ideally, *ARSpectator* should be used through an AR HMD, presumably thin and light for improved comfort. Spectators wearing the AR HMD then would need to spectate a game like they usually do, and the system works behind the scenes to present useful situated visualization to the users without requiring much effort from the user. This would be an attempt toward an invisible user interface (Weiser 1998), where the system understands what the user wants by studying the spectator's behavior.

In the stadium, spectators need to interact with their environment by looking at what they are interested in. For example, if the spectator is interested in knowing more about a particular team, spectators would need to focus on the team's score in the stands, and situated visualization would appear regarding the team. Based on the context, other visualizations will also appear. Some questions would be asked before using the application to determine the use case and knowledge level of the spectator. A novice spectator would likely be interested in the explanation of every event. In contrast, an expert spectator would most likely want some statistics that the spectator cannot obtain simply by watching the game. We discuss details about user personalization in Chapter 6 of the thesis.

This thesis only details the visualization and user experience of the research project. The overall *ARSpectator* research is still ongoing. Although this involves many different parts of the project done by other researchers, it is helpful to know the overall system to understand better how everything works. Therefore, we start with an overview of *ARSpectator* to explore the interactions between each component.

3.1 *ARSpectator* Overview

Usage of AR in a large-scale environment is often challenging because of the localization and alignment of the users' devices with the environment. We need to identify the position and orientation of the user to provide properly aligned graphical content. These were less of an issue in a traditional small work-space AR (Klein & Murray 2007), in which the device circles around the object of interest. In our case, any misalignment would be more evident due to the distance of the visualization from the user. The users' position, however, usually is static while the action happens around

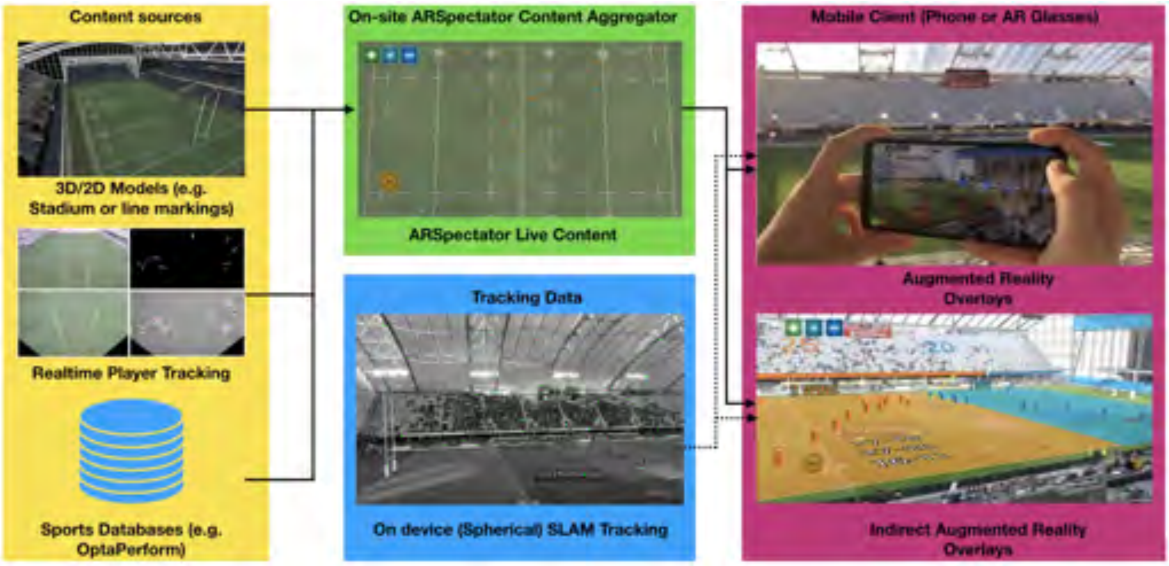


Figure 3.1: Version 1 of the *ARSpectator* overview diagram. *ARSpectator* combines sports data and player tracking to provide spectators with situated visualizations on their mobile devices (mobile phones or AR HMD). Figure from (Zollmann et al. 2019).

the user. This has reduced the complexity of tracking as we are less concerned with spectators moving around to different positions in the stadium.

3.1.1 Components of *ARSpectator*

Multiple components work together to provide users with what they need to see. In general, the *ARSpectator* system consists of four main components —the **content sources**, the users' **tracking data**, and the **content aggregator** which leads to the **mobile client** where the visualization shows (Figure 3.1). With the progression of this research, we ended up with an updated version 2 *ARSpectator* overview diagram (Figure 3.2), which provides more details on the interaction between the sub-components in the four main components mentioned above. We designed the overview in a way in which the components range from the server-side implementation (left side) to the on-device processing and visualizations (right). In the middle, there is a mixture of server-side implementation and on-device processing in the content aggregator and user movement tracking. Of course, many other components exist, such as the stadium

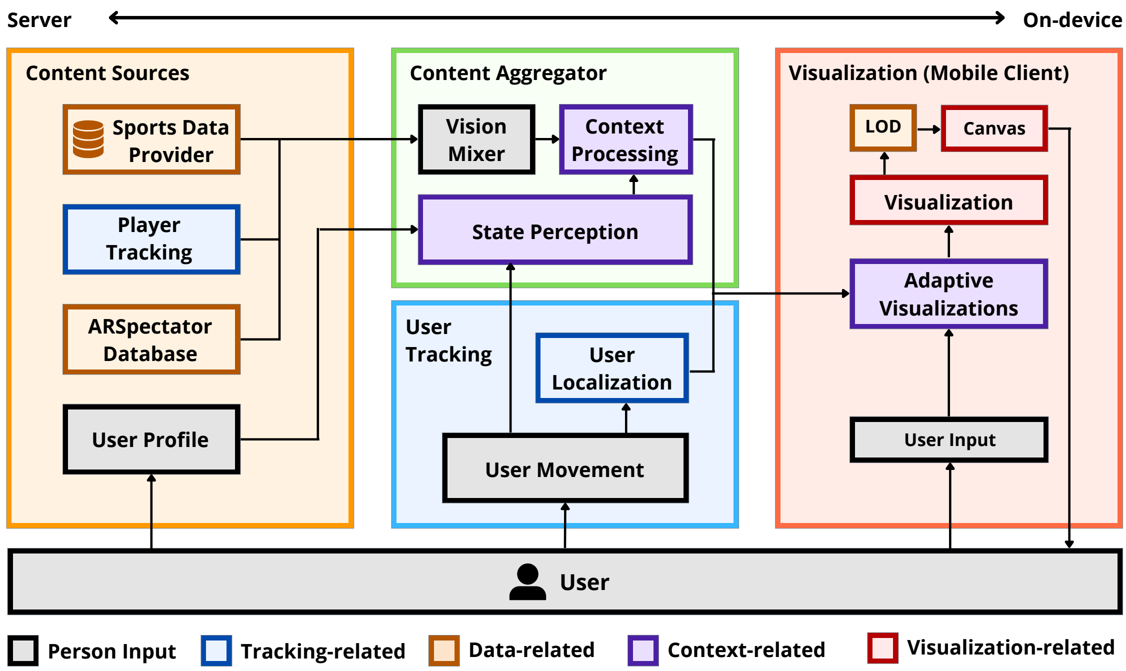


Figure 3.2: The updated version 2 *ARSpectator* overview diagram. This new overview shows in more details the various components involved in the *ARSpectator* system and how the user is involved.

infrastructure and the players, that facilitate the *ARSpectator* system as a whole. However, we omit those, as we are just looking at the technical components of *ARSpectator*.

Content Sources

The content sources of *ARSpectator* come mainly from the content server. The *ARSpectator* content server is a computing device that is connected to the **sports data provider**, **player tracking cameras** and ***ARSpectator* database**. The content server manages these content sources and sends them to the content aggregator for further processing. Another input source is the **user profile**, which allows the system to compute what the user is interested in seeing. Although we have a hosted content server, most of our user studies are conducted with local content on the mobile device to reduce confounding variables and complexity. This approach eliminates the server's latency in all of our user studies.

Sports data providers are external companies that provide real-time sports data and statistics, such as OptaPerform[1] in our case. There are many types of data, ranging from field coordinates to the type of events and players involved, where the data will be combined with the information from the player tracking. The tracking of the players could be done through hardware tracking such as sensors embedded in the jerseys of the players (Kirkup et al. 2013, Blauberger et al. 2021) or by using computer vision with the help of cameras (Baysal & Duygulu 2015)).

The *ARSpectator* database is a database of information related to the game, venue, and everything else. One of the main elements we have in the *ARSpectator* database is the 3D model of the stadium. The professionally surveyed 3D model is created by ARL[2]. It is used during development to attach visualizations into the right positions based on the actual environment. We could obtain user profiles in two ways: questionnaires asked before the game or through pre-existing profiles. User profiles could potentially include past user behaviors to predict better what the spectator wants to see during a game.

Content Aggregator

The content aggregator is similar to the "brain" of the system, taking input from various sources and compiling them into something comprehensible. It then decides when and where to display this information through visualizations. The content aggregator includes the **vision mixer**, an optional component which is an on-site person who does annotations on the fly as the game progresses. The vision mixer plays the same role as the broadcast operator for live sports broadcasts, where the operator would annotate or select different visualizations for spectators to see. The vision mixer is optional as the *ARSpectator* system does feature visualizations created in advanced and could use game context to determine what visualization to show. Despite that, an experience vision mixer would be faster in determining what is happening and could predict what visualizations might be needed in certain situations. When the components complement each other, this could provide additional input to the system and potentially improve the output's performance.

[1]https://optasports.com

[2]Animation Research Limited. URL: https://arl.co.nz/

The content aggregator also consists of the **state perception** component that tries to identify the spectator's state during the experience, such as is the spectator looking for information or are they focusing on a single player etc. Taking into account the movement of the user, the state perception component then combines this input with the user profile and tries to predict what the user would be interested in. We then parse these data into the **context processing** component, where it takes the output of the vision mixer (if there is one) and combine it with the spectators' state to determine which visualization to show and how detailed should the visualizations be. If there is no vision mixer, the context processor will utilize the game state alongside any other supplementary information and try to play the vision mixer's role. Chapter 6 will document the different states we identify in both the game and the spectators and how we developed an interface for *ARSpectator*.

User Tracking

Since the users get a tailored perspective of the visualizations, **user movement** is vital to provide details on what the user is looking at. The user movement is also the input for the state perception mentioned in the previous paragraph. The accelerometer, gyroscope and magnetometer on the device determines the user movement, taking the three axis pose of the device to calculate the users' pose (Desai et al. 2014). Depending on where users are looking or how quickly they turn their heads/devices, we can roughly estimate where a user is looking and therefore predict intentions depending on where users are looking.

The other important component is **user localization**, where we need to determine the position and orientation of the AR client in the stadium. The system first needs to know where the user is located in the stadium environment and then continuously track the user's pose to provide situated visualizations tailored to the user's perspective. We implemented two methods of user localization, the first being a *user-guided localization* and the second, an *automated localization*. The following two paragraphs use content from Zollmann et al. (2019) and Zollmann et al. (2022) to explain our approach to user localization.

The user-guided approach uses the spectators' seats to approximate spectators' position in the stadium. Using the estimates, we generate an AR overlay of the field, which the spectator then tries to align with the actual field. Some rotational controls were provided for spectators to get the best fit. The challenge for this method is that we would need to have a spatial map of all the stadium seats, which might not be available. We also implemented a traditional perspective-n-point method that allows spectators to select 4 points of a field. However, due to the narrow FOV of the devices, it is not always possible to get all the corners in view.

For the automated approach, we implemented a vision-based approach, using stadium advertisements as image targets. We did obtain good results for certain advertisements, but performance with image targets on the field is highly unstable. This method, however, depends on the advertisements' placement, which changes from time to time and is also highly dependent on the spectator's position in the stadium. Since we have our stadium CAD model, we also did model-based implementations where the system tries to match the camera view to the stadium model. This approach is also location-dependent and hard to achieve due to the repetitive nature of the stadium structure. (Baker et al. 2018, Zollmann et al. 2019)

Visualization (Mobile Client)

The mobile AR clients are the devices used by the spectators on-site that serve as input and output devices. They could be in the form of a smartphone or, in an ideal situation, an AR HMD. We believe that AR-enabled smartphones are a good temporary solution as AR HMDs are not yet accessible in terms of portability and pricing. Mobile AR still provides the information spectators would want at the sacrifice of ergonomics. This is because spectators will experience arm fatigue if they hold a smartphone for a while. Re-initializing the device in the stadium would need to be quick to avoid distracting the spectator and missing any action.

The mobile client is where the visualizations are presented based on the context from the content aggregator and tracking components. Through the mobile AR client, spectators would then be able to see visualizations that originate from the sports data provider, undergo some context processing, and end up aligning well in the spectators' FOV of what they want to see. It starts with the **adaptive visualization** components that compile the output of the context processing component and the user localization. Then it determines which **visualization** to show, in what **level of detail (LOD)** before finally deciding on which **canvas** it should be visualized on. The process then ends with the system presenting the visualization to the user. Chapter 4 discusses the different visualization methods developed and the framework we use to create the situated visualizations.

3.1.2 *ARSpectator* Challenges

While attempting to integrate AR into live sports events, we encountered a few hurdles to be solved before we could materialize an on-site AR sports spectating application. Although it does not relate entirely to the visualization and user experience of *ARSpectator*, we would like to briefly describe some of the challenges we face in the development of *ARSpectator*.

Latency: This is probably one of the biggest challenges we face, considering that there are delays even for live sports broadcasting. Most visual overlays in sports broadcasting are also not computed in real-time; hence, the latency in the visualization pipeline will not be as critical. However, since our application is real-time, latency is an essential factor to consider. In our prototypes, depending on the infrastructure, we experience practical latency above 500ms (Zollmann et al. 2019).

Bandwidth: While this is not an issue for the development and testing process of *ARSpectator*, it is something to consider in a real-world use case of *ARSpectator*. A low-latency network with a large bandwidth would be required in the stadium to have thousands of devices simultaneously connected to the internet and receiving real-time data from the content server. 5G technology could answer the problem as the emergence of 5G-enabled venues increases (Wu et al. 2022), but this is not in our research scope.

Localization: Although we presented some localization approaches, none of the approaches is a perfect fit for our problem. They all work in certain conditions (depending on spectators' seat position, FOV, stadium conditions, etc.), and we think there is still room for improvements to find a robust approach.

Visualization: There are various visualizations done on sports broadcasts that we could take as inspiration. However, most of them are catered for sports broadcasts and are viewed from a single perspective, the camera. Not many sports statistic visualizations also use spatial relevance, such as placing certain graphical content on the pitch for increased understanding. Therefore, many of our visualizations would need to be designed based on the canvases and scenes.

Usability: While *ARSpectator* is planned for AR HMD, this technology is currently not accessible in terms of price point and portability. Using *ARSpectator* on a mobile phone introduces arm fatigue, which participants in our user study encountered in Chapter 5. Therefore, *ARSpectator*, in its current state, still lacks usability.

On-site Accessibility: Last but not least, we have the issue with on-site accessibility. Access to the stadium is already limited due to security and logistic issues. Therefore, it was not easy to develop an on-site application while off-site. The situation only worsened with the COVID-19 pandemic, which further restricted access to the stadium. For this challenge, we created multiple prototypes and developed a flexible XR prototyping framework that will be discussed in the next section.

3.2 A *Flexible XR Prototyping Framework*

Over time, we found that other prototypes in different modes are needed. The main purpose of having flexible XR prototyping is to overcome hurdles regarding the accessibility of the on-site environment. Even research for somewhat accessible locations (such as a park) is often done in a different location, possibly due to financial and time reasons. The *Flexible XR Prototyping Framework* (Figure 3.3) is built on the implementation of our AR sports spectating use case in which users visit a stadium environment and receive visually overlaid situated visualization of the sports event that is happening. In our main scenario case, it is a game of rugby. However, our objective is to generalize to other AR application scenarios by introducing different characteristics, prototypes, and components for seamless development and evaluation experience. In this section, we discuss the framework starting from the considerations,

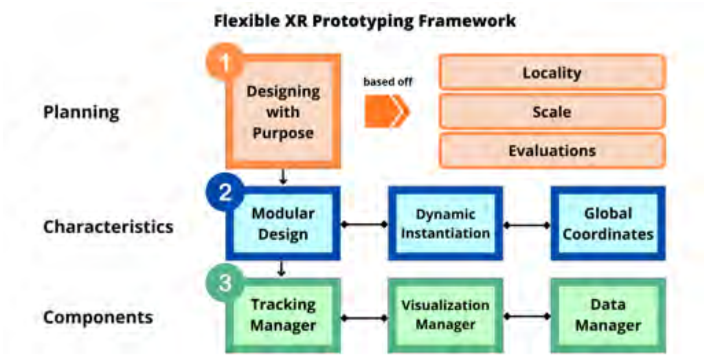


Figure 3.3: Flexible XR Prototyping framework: This framework guides researchers to have proper planning alongside the required characteristics and components for a flexible XR prototyping development process.

the characteristics, and the components needed. In the next section, we explain the implementation and distinct differences between the various prototypes in more detail.

3.2.1 Designing with Purpose

"What prototypes do we need to develop a mixed reality application" is not a simple, straightforward question to answer, as most prototypes are born out of necessity later. The *Flexible XR Prototyping Framework* prepares researchers to create new prototypes without producing too many complications and interruptions at later stages of the development process. However, it is good practice to brainstorm what we potentially need throughout the development process to make early design choices with better judgment and knowledge. Based on our own prior research and development experience with AR and XR prototypes while developing ARSpectator, we focus on the following aspects.

Locality

Locality refers to where we intend to use the prototype. It is probably one of the most determining factors in deciding whether one's research needs to incorporate a particular prototype. Researchers need to question themselves if they need an off-site prototype. An off-site prototype is necessary for large-scale environment AR research (White & Feiner 2009*a*, Grubert et al. 2016) such as AR for city navigation, where the subject of interest is not something on a tabletop like labeling of an object (Kalkofen et al. 2013, ElSayed et al. 2016*a*). Justification for this could be ease of development, evaluation, off-site demonstration, etc.

We have always had challenges accessing the stadium due to security and logistics issues in terms of paper-works and applications, which became increasingly harder due to the COVID-19 pandemic[3] mid-way through the thesis research. Therefore, we developed a mobile indirect AR and a VR indirect AR prototype, also known as our VR prototype. As the names imply, the mobile indirect AR uses a mobile phone as the hardware, while the VR version uses a VR HMD. These are suitable for remotely experiencing large-environment AR where users could be immersed in their environment while having control of their viewpoints and viewing angles. However, we also used a scaled-down prototype to experience a large-environment prototype which we discuss in the next section.

Scale

The scale here refers to the scale of the environment in which the XR prototype will be used. Using printed image targets, we recreated a smaller scale of the actual stadium environment with the same visualizations we call *miniature lab AR*. This prototype can be imagined as having the stadium in AR right on the surface of a table, where one can walk around, seeing the visualizations from a bird-eye view. Often, it is easier to get an overall picture of how visualizations are performing when seen from a third-person Point-of-view (POV). For instance, developers could target the question of whether there are any occlusions between visualizations that might not be noticeable from the first-person POV. The miniature-scaled AR prototype also aids in evaluating visualizations since the user is free to maneuver to different spots in the AR environment relatively unconstrained. Developers can even access some spots that might not be easily accessible from a first-person POV.

[3]The restrictions in New Zealand due to the COVID-19 pandemic meant we could not do on-site testings during lockdowns.

This miniature lab AR prototype also benefits from off-site demonstration because it is more portable due to the smaller scale. There were many occasions where we used the miniature lab AR in demonstrations for interested parties, lab demonstrations, and conferences. We just brought along a foldable canvas pitch image target and a mobile device with our prototype installed. This prototype also helped the general public understand what AR is about without having to demonstrate it on-site at the stadium.

Evaluation

Evaluation refers to both the researcher's perspective and the user study perspective. As mentioned in the section 3.2.1, if a prototype is designed to work in a specific location, how do researchers remotely demonstrate it, such as at a board meeting or a conference? Researchers may also face problems running on-site user studies if there are confounding variables such as noise, distractions, and uncontrollable events. These confounding variables are important in a final usability testing environment but could weaken the integrity of the evaluation if it is only focused on a specific area. In our user studies, we used the mobile indirect AR prototype to standardize the visualizations as we wanted to avoid ambiguity from random events.

For some AR research, the locality might involve some element of risk, such as using AR in vehicles. A more controlled environment or a simulation might be an appropriate alternative. These issues could be resolved or reduced with an indirect AR prototype. We could not conduct our preliminary user studies at an actual game for the sports spectating use case. This was not only due to confounding variables but also the logistical challenge of distributing our devices as the application was not published. Also, monitoring participants in a crowded stadium is hard, especially when alcohol is served on-site as they might lose focus in the study or get disturbed by other drunk spectators.

3.2.2 Characteristics

We will now consider some of the characteristics that these different prototype classes should have in order to ease the development process.

Modular Design

A modular design allows scalability while streamlining the development process when multiple prototypes are present. Researchers would need to think about the development of prototypes in the form of modules, in our case, the tracking, visualization, and data manager, which we will describe in the components subsection. Before adding other modules, all prototype developments probably start with some base module for minimal recreation of the environment.

The idea with these modules is that they will each share the same code base or prefabs, as they are called in-game engines. Therefore, changes made to the visualization module of the indirect AR prototype should appear in all prototypes, regardless of which prototype we apply the changes to. The benefit of this approach is that changes will be available across all prototypes, saving the developer time to implement them one by one. This will decrease the chances of error and inconsistency. However, it also carries the risk that a modification made for prototype A might indirectly affect prototype B without the developer noticing. Hence, more frequent testing of other prototypes is required, but when done correctly, it promotes good coding practices and easier debugging while saving more time in the long run with more efficient code.

Dynamic Instantiation

Dealing with multiple prototypes often means developing in multiple scenes, even if it is in one project. The manual way of doing it is to individually create and place every object and visualization in the scene, as we are currently doing with our sports spectating use case. This process is time-consuming and would result in errors if the objects in the scenes are not synchronized. Therefore, it would be very beneficial to the development process if we could automatically instantiate most of the dynamic content in the AR experience by script. In this way, developers would need to test out the visualizations in the editor and then instantiate the visualizations as prefabs via scripts. It is slightly more work in the beginning but would greatly benefit the development process if there are multiple prototypes where visualizations or game objects are shared.

Global Coordinates

As AR research usually involves the definition and use of many different coordinate systems, all coordinates should refer to one global coordinate system, which is visible and measurable in the real world. We used one corner of the playing field as our point of origin in the stadium; therefore, in all of our prototypes, the position vector (0,0,0) would point to the same spot. This standardization greatly assists in scenarios where there is object tracking data as an input source, in our case, player tracking data and event-based data. With this approach, all appropriate visualizations appear at the same position for all prototypes, reducing the trouble of individually translating incoming vectors to suit each prototype's coordinate spaces.

3.2.3 Components

The framework components are some of the crucial modules mentioned in the characteristics. Almost every AR application would have these three essential aspects — tracking data, visualizations, and incoming data sources (Zollmann et al. 2020). In the AR sports spectating use case, we have three main modular components, which are the tracking, visualization, and data manager; where in our implementation, each of the managers is a prefab game object in Unity consisting of scripts shared among all prototypes.

Tracking Manager

The tracking manager manages most of the tracking and localization performed in the application. Therefore, the tracking manager exists only in our stadium AR and lab AR prototypes. In our case, the tracking manager consists of two sub-components, one being the *image target manager* which deals with the few image targets we have scattered throughout the stadium environment, and a *manual target registration* which stores the details of various seat positioning and manual controls for initializing the stadium in the stadium AR application.

Data Manager

During a game, we obtain data from various sources, from the sports statistics provider to data from other mobile devices for crowd-based interactions and engagement. The data manager is the module that handles these incoming data and sends them to the correct destination to be visualized or processed. It takes in data controllers that process raw data from the sports statistic provider, usually in the form of XML queries, so the coordinates and details can be correctly translated.

Visualization Manager

The visualization manager is the most extensive module, as it contains many scripts related to visualizations — from the actual stadium model itself to the text augmented on a spectator stand. The visualization manager takes the output from the data manager and feeds it to the different attached visualization scripts to show the correct visualization at the right time. The visualization manager also contains some game objects such as colliders that help detect if a user's gaze is colliding with that particular collider, providing insights on where the user is looking. Everything is manually added to the manager when a new visualization is created, but we look into automating it in the future as part of the dynamic instantiation characteristic.

3.3 Prototypes — Framework Implementation

We did bring up a few of our prototypes developed during the introduction of the *Flexible XR Prototyping Framework*. We developed our prototypes with a modular design, and we used global coordinates; however, dynamic instantiation, which we derived after development, is planned for future work. We describe the different prototypes that we developed and the challenges that we faced. For our project, we developed four prototypes, mainly mixed fidelity, each with its use-cases (Figure 3.4).

We created Figure 3.5 to show the main differences between the four XR prototypes we developed. The stadium AR and lab AR prototypes use visual-inertial odometry (Gui et al. 2015) to track users' position, using a combination of cameras and on-

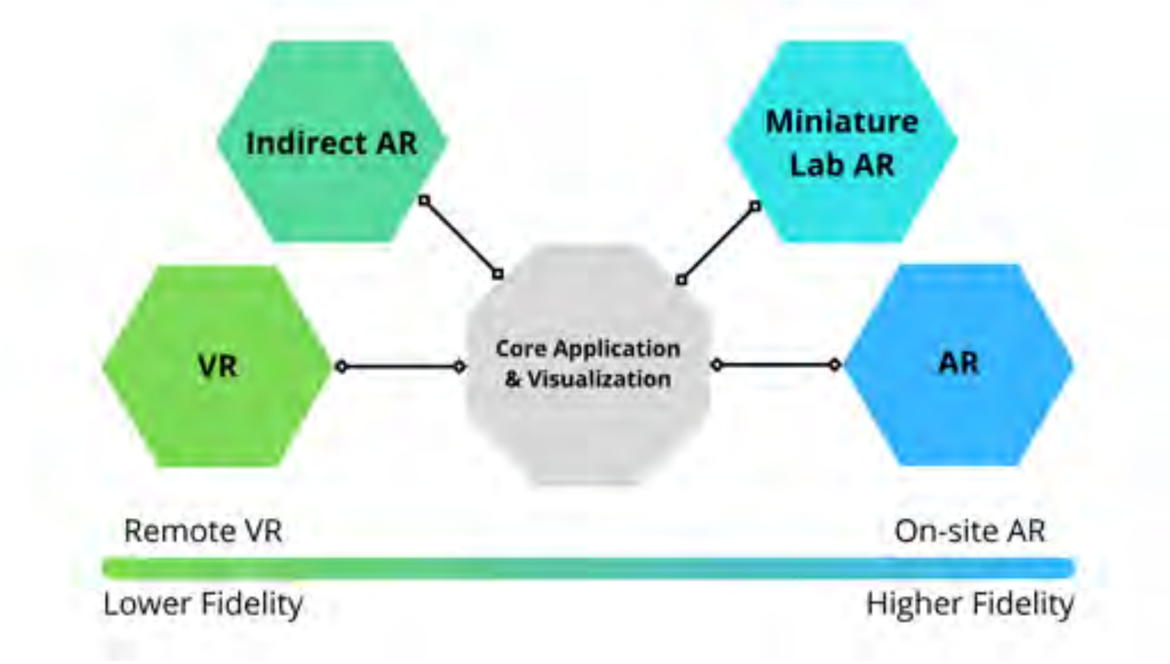


Figure 3.4: Overview of our sports spectator XR prototypes, all originating from a core application where main components are synchronised between each prototypes.

	Prototypes	Live view of the actual environment	Simulated Stadium Environment (by 360° video)	Implemented Use Cases	Tracking Methods (Implemented)
On-site AR ↑	Stadium AR Prototype (Mobile/AR HMD)	✓		Assist on-site Viewing	Image Target Seat Localization Visual-inertia odometry
	Lab AR Prototype (Mobile)	✓		Off-site R&D assistance* Table-top remote viewing	Image Target Visual-inertia odometry
	Indirect AR Prototype (Mobile)		✓	Off-site R&D assistance*	On-device sensors
↓ Remote VR	VR Prototype (VR HMD)		✓	Off-site R&D assistance* Remote viewing	On-device sensors

*Off-site R&D assistance refer to research, development, user evaluation and demonstration out of the stadium

Figure 3.5: A table showing the differences between the XR prototypes. The prototypes are sorted from an on-site AR application to a fully VR prototype that could be used anywhere.

device sensors. All prototypes are created with the Unity game engine (versions 2019 and 2020) and contained in the same project within different scenes. We used Vuforia[4] for image target tracking and extended tracking in prototypes that needed them. The basis of all prototypes was the same: a stadium environment with visualizations attached to it. However, the differences are in the implementation, use case, scale, and environment in which the prototypes are being used. To place the augmented content in a spatially correct way, we utilize a 3D CAD model of the stadium to assist in anchoring visualizations. The model aids in the process of locating canvases in which we can place visualizations that are coherent with the environment. Visualizations in the actual AR application consider the position and viewing direction of the user in the stadium.

3.3.1 On-site Stadium AR Prototype

The on-site stadium AR prototype (Figure 3.6) is the closest representation of our final product that will be used in a mobile device. The prototype serves as a tool to provide additional information about the game in the form of situated visualizations during a live game. This prototype is done as a mobile AR application. We experimented with the Microsoft Hololens AR HMD, but the tracking was not as good as the mobile AR version. This is due to the Hololens being catered towards using image target registration, in which it was relatively hard to initialize in the stadium as we do not have a good view any image targets.

Unlike some of the other prototypes that do not have a user localization component, the AR prototype uses all the components mentioned in the *ARSpectator* overview (Figure 3.2). The system needs to localize and track the users' movement to present well-aligned visualizations in the actual environment. Data arriving from the content sources would also need minimal delay due to the real-time nature of this prototype. Hence, this is also the most complex prototype to develop among the four. Due to the pandemic, this prototype was also the least used due to the lesser opportunity we had to conduct tests in the stadium and the tracking challenges it faces.

[4]https://www.ptc.com/en/products/vuforia

Figure 3.6: On-site Stadium AR Prototype closely resembling the final product. Players on field are real players.

The challenge of an AR prototype on-site, as mentioned, was accessibility to the stadium venue. Besides that, since spectator movement is limited during a game, it is hard to initialize the localization process using an image target. Depending on seating, the image target might not be visible or viewed from a perspective the camera could not recognize. Therefore, another approach to tracking and localization is needed for the on-site prototype. We looked into different solutions such as spherical Simultaneous Localization and Mapping (SLAM) (Zollmann et al. 2019) and line homography (Skinner et al. 2018) to try to localize the users in the stadium, but that is not in the scope of this thesis.

3.3.2 Miniature lab AR

The miniature lab AR is a small-scale version of the final product, designed to be used in the laboratory or in situations where portability is needed (Figure 3.7). Utilizing

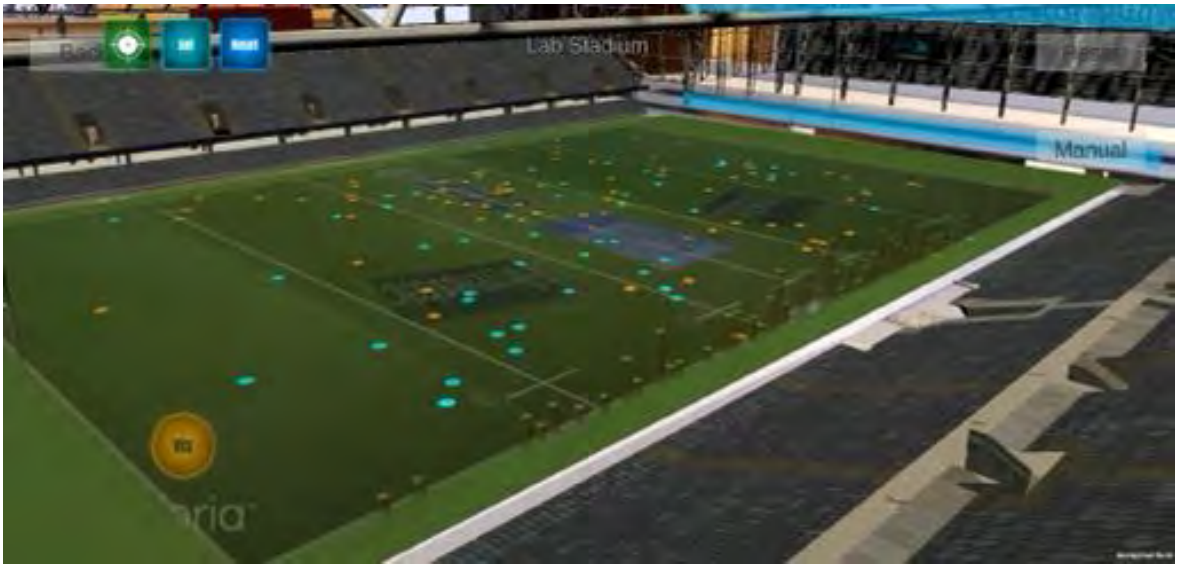


Figure 3.7: Miniature lab AR prototype to evaluate visualizations from a third-person perspective. Shown is a virtual stadium overlaid on a A0-sized printed field.

a big A0-sized printed field with advertisement logos as an image target allows for a bird-eye view of the stadium model while still allowing the various visualizations to be shown. Due to the smaller scale, this prototype is the only prototype that allows the user to have a "God mode" where they can walk around the stadium and view visualizations from different perspectives, including moving through the structure and viewing visualizations from inside the stadium.

This prototype also contains all the components mentioned in the framework since it does have tracking, albeit on a different scale. Tracking is easier in this prototype as the image target on the field are more visible. Since, even in the stadium, we cannot simply move around quickly to the opposite stands to test our visualizations, this prototype serves as an ecologically valid alternative by allowing evaluation of visualizations from various perspectives to, e.g. prevent occlusions. All of these can be done right in the lab without needing to visit the venue with a more straightforward interaction method than manipulating a CAD model on the computer. On certain occasions, we used the miniature lab prototype for user studies in a controlled lab environment. We investigated spatial understanding in situated infographics compared to traditional

Figure 3.8: Indirect AR Prototype mimicking what spectators would see at the stadium through their devices.

on-screen infographics.

This prototype is more suitable for a single-person experience or small groups where they can move around the printed poster. It is hard to demonstrate to a large crowd depending on where the person is in the environment, and the perspective would be very different from the one in the lab AR situation. In addition to that, occasional tracking errors cause misalignment of the stadium. Apart from that, since visualizations appear smaller in the lab AR prototype due to the scale, it might give a false impression of the visualization size and orientation compared to the actual use-case in the stadium. Often, visualizations look just right on the lab prototype but are too large in the actual stadium-scale environment.

3.3.3 Mobile Indirect AR

The mobile indirect AR prototype (Wither et al. 2011) is the leading prototype used for demonstrations outside of the laboratory and user studies (Figure 3.8). We started using 360° panoramic photos (instead of videos) captured in the stadium via a Ricoh Theta S[5] to simulate the spectators' viewpoint, as the Theta S only takes low-

[5]Ricoh Theta S, https://theta360.com/en/about/theta/s.html

resolution video. Upon upgrading to an Insta360 One X[6], we replaced still images with 360° videos. This gives users the freedom to look around the virtual environment where we place the situated visualizations. Depending on the scenario, different types of 360° videos have been used, such as an empty field or an actual game. This prototype proved to be very useful in the development and testing of visualizations as it is independent of the tracking challenges faced during an on-site testing environment or with the miniature lab AR prototype, meaning that it does not have the tracking manager from the framework.

Since we do most of the prototype demonstrations out of the stadium, this prototype resembles using a mobile phone on-site. Users could see the visualizations as if they were sitting in the stadium and could do almost everything that could be done with the actual AR prototype. However, this prototype does not have the user localization component as the user is technically placed in a fixed position. This makes it the best prototype for a user study evaluating visualizations in mobile AR since there will be fewer confounding variable effects regarding tracking and localization. Despite being in the stadium, we conducted one on-site user study with the mobile indirect AR prototype. Participants viewed a recorded game while physically in an empty stadium to get a better understanding and immersion with the sports spectating use case. The mobile indirect AR is also more predictable and reliable in data visualization compared to obtaining information from a live game.

One of the limitations of the mobile indirect AR is the quality of the 360° videos since it is lower quality than their mobile devices' cameras. This is due to the technical difficulties of recording a 360° video at high resolution. Despite us recording the video with an Insta One X at 5.7k resolution, the footage is relatively blurry compared to a standard mobile device's camera as the 5.7k pixels are spread out in a sphere. We are still trying to overcome this issue by using a professional-grade 360° camera such as the Insta360 Pro 2, with multiple small cameras recording simultaneously. However, this device costs significantly more than the consumer-grade 360° cameras and would introduce a huge video file, which might not be feasible for this prototype.

[6]Insta360 One X, https://www.insta360.com/product/insta360-onex/

3.3.4 VR Prototype

The VR prototype continues the mobile indirect AR prototype but is used in a VR headset. This prototype aims to recreate the closest experience to using an AR HMD in the stadium without needing to localize the user. By using the indirect AR prototype in a VR headset, users can turn their heads around to look at the stadium surrounding while spectating a pre-recorded 360° video of a game alongside situated visualizations. Due to the lack of a reachable touch screen, this prototype forces the testing of alternative interaction methods as would happen when using an AR HMD, where users cannot interact via touch screens. This prototype also retains the advantage of the mobile indirect AR prototype, which eliminates tracking and localization issues compared to the other prototypes.

We use this prototype to test a "hands-free" interaction experience as participants often complained of arm fatigue when using the mobile indirect AR for a prolonged period. Our prototype allows us to research the gaze-based input at the center of the screen. Thus, we can collect better data on where the spectators are looking during the experience. From our observations with the mobile indirect AR, some participants did not move the mobile devices much; presumably, this concept of AR is still relatively new to them. On the contrary, in VR, moving one's head to view something is more natural and is easily picked up by participants.

Although our prototype might seem like the best off-site prototype to be used, it still comes with its shortcomings. The effect of lower resolution 360° video viewing is slightly amplified in VR, where the pixel density is more spread out than on a mobile phone, making the quality difference even more noticeable. The other issue involves user studies in which facilitators of the user study would not be able to see what is happening in the VR HMD entirely. One way of addressing this challenge is to use casting options such as the Oculus Cast. However, the FOV in such options still differ from what the actual participant sees in the headset.

3.4 Summary

This chapter describes the overview of *ARSpectator* and the process undertaken to include other XR prototypes. Due to this, we developed the *Flexible XR Prototyping Framework* to create standardization across different prototypes. We then provided four different prototypes developed during the *ARSpectator* research as examples to support our proposed framework. Each prototype plays a specific role in the development process of *ARSpectator* and is very useful during evaluations and demonstrations.

Chapter 4

Situated Visualization for ARSpectator

Contents

So far, this thesis has documented the motivation for using AR in on-site sports spectating and provided an overview of the various components. The development process was explained along with the needs of different prototypes. This chapter focuses on the visualization aspect of *ARSpectator*. We will explain the proposed *3'C's Situated Visualization for on-site Sports Spectating Framework*, consisting of the three main components — canvas, content and context. This determines where and when to display what kind of information. We will also introduce two different visualization methods we developed for our prototypes and discuss the pros and cons of each method.

As highlighted in Chapter 1, situated visualization seems to be a promising solution for using AR in sports spectating. However, to date, it remains unclear whether spectators would benefit from such on-site visualizations and the best way to implement them. In this chapter, we address this gap by proposing a conceptual framework that describes components of situated visualization for on-site sports spectating. Although our scenario mainly focuses on rugby (Rugby Union), we have designed the framework to apply to most sports, especially team sports. We used the *3'C's Situated Visualization for on-site Sports Spectating Framework* to develop and evaluate two situated visualization approaches — *Situated Broadcast-styled Visualization (SBV)* and *Situated Infographics (SI)*.

In order to evaluate our situated visualization methods, we conducted user studies in a controlled lab environment and on-site in a stadium. Our findings show positive responses towards the concept of on-site *SI* and an improved understanding of the game without an increase in mental demand compared to on-site spectating without assistance or with traditional infographics (TI). The reported user experience of the proposed visualization methods was also positive. However, we can make improvements to make it more inclusive for spectators of all knowledge groups, which we discuss in Chapter 6 as part of context-awareness. To replicate the live event experience, we also explored the use of indirect AR (Wither et al. 2011) for the sports spectating scenario, for both testing and actual on-site use.

4.1 A Conceptual Situated Visualization Framework

The main goal of developing situated visualization for on-site sports spectating is to elevate user experience through a better understanding of the game. However, in addition to creating sports statistics, we emphasize specific considerations to make spatially relevant visualizations. Previous frameworks (White & Feiner 2009*b*, Bach, Sicat, Pfister & Quigley 2017) are often too general and lack in sports contexts, such as different types and temporal components of sports data, the users' involvement as a spectator, and identifying where content should be placed. Therefore, based on previous work, we conceptualize the three components, which we dub the *Three 'C's* that are essential pillars for developing AR situated visualizations in sports spectating. The three components are ***canvas***, ***content*** and ***context***. We will discuss how we

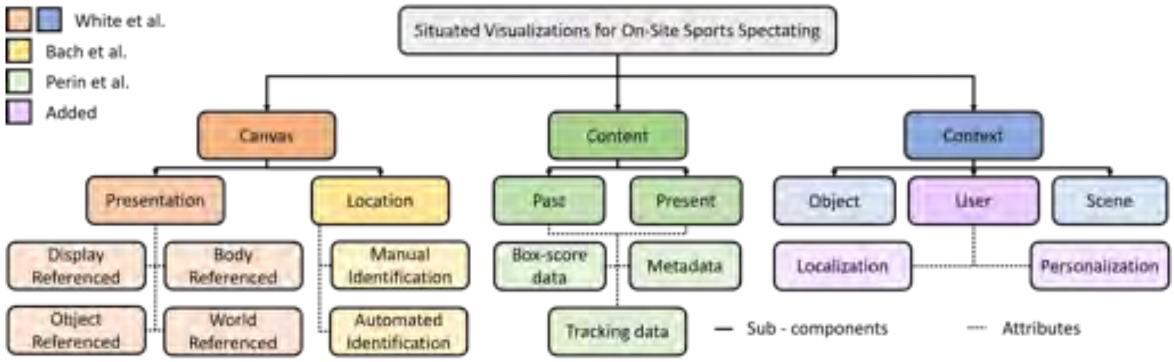


Figure 4.1: Our proposed *3'C's Situated Visualization for on-site Sports Spectating Framework* depicting the three main components: *Canvas*, *Content*, and *Context*. We highlighted the adapted components from previous work (White & Feiner 2009*b*, Bach, Sicat, Pfister & Quigley 2017, Perin et al. 2018) as well as new components with color codings.

formed these components and their relevance. In the following, we will refer to White et al.'s visualization framework (White & Feiner 2009*b*) as FW1 and the AR-CANVAS framework (Bach, Sicat, Pfister & Quigley 2017) as FW2.

4.1.1 The Three 'C's

To better fit our sports spectating scenario, we decided to build on FW1 to incorporate elements of sports spectating while taking inspiration from other works such as FW2 and the *5'W's and 1'H' user-centric concept* (who, where, what, when, why, and how) (Jang et al. 2005). The three components dictate where a visualization should go (*canvas*), when it should appear (*context*), what it should visualize (*content*), and why we visualize something. Under each component, there are sub-components and attributes which allow us to further specialize in different AR sports visualization approaches. These main components are crucial and are present in all cases of situated visualizations in sports spectating, while the sub-components and detailed attributes are optional. We will discuss the components of the framework (Figure 4.1) in detail in the following.

Canvas

The *canvas* is a dynamically assigned and positioned plane or surface where we could anchor visualizations to provide spatial relationships to otherwise non-spatial data. FW1 discussed the concept of the *locus of presentation* while FW2 discussed the location of the *canvases*. In our framework, we bring these two sub-components from FW1 and FW2 together, integrating the different types of presentations with either a manual or an automated approach to canvas identification.

Concerning FW1, the presentation of the data consists of display, body, object, and world-referenced visualizations. It covers the spectrum from screen-based visualization to scene-based visualization. Display-referenced visualizations are visualizations that take place on the screen space itself, regardless of where the user is looking. Body and object-referenced visualizations are anchored to an object; in the case of body-referenced, the users themselves. World-referenced visualization uses the surroundings as an anchor, not necessarily attached to a specific object.

Events where situated visualizations for sports spectating are relevant take place in known environments such as a stadium environment. 3D models of the environments are often available for large venues like this. In such cases, developers and designers can apply a manual approach to identifying static *canvases* and can define canvas options using such a model of the venue. Manual approaches, while usually taking more resources, provide better results than automatic approaches because there is a lower rate of placement errors and the visualization is more custom-fitted. Automatic approaches include plane detection methods that identify flat surfaces in the physical environment (Nuernberger et al. 2016), image space analysis (Langlotz et al. 2014) as well as using geographic information systems (GIS) (Skinner et al. 2018). GIS data can be combined with image analysis to obtain more accurate representations of a structure and make it better suited for real-time implementation (Zollmann & Reitmayr 2012).

Content

Once we have identified the *canvas*, we need to prepare the content we want to display on those *canvases* before placing the visualization. FW2 classified all *content* into a category of general context data, while FW1 got inspiration from the basic classifications of nominal, ordinal, and quantitative data (Card & Mackinlay 1997). Since our framework focuses specifically on sports visualization, we based our *content* component on the classification of sports data by Perin et al. — *box-score data, tracking data* and *meta-data* (Perin et al. 2018). We also have a categorization of past and present data to emphasize the temporal properties of the data.

Box-score data is the most common form of sports data. Originating from the box-like format of score-taking on paper, it features the recording of discrete events in the game, such as scores, fouls, substitutions, etc. Tracking data involves more dynamic data, such as player positioning and ball placements, allowing the generation of *dynamic canvases* and the creation of embedded real-time visualizations, often used for sports performance analysis (Perin et al. 2013, Gudmundsson & Horton 2017). Lastly, meta-data comprises all other data that is not directly related to the game, from historical facts such as previous match-ups and venue history to dynamic data such as crowd emotions and engagements.

The type of information matters in our framework because one could present visualizations in different formats, times, and places. Traditionally, box score and tracking data are shown when the game is ongoing, requiring real-time processing. Tracking data by itself requires auto-generated visualizations as there is no time for broadcast operators to annotate the visualization unless it is a replay of past events. Some meta-data such as information regarding game venue, previous match-ups, and weather details would be shown mainly during pre-game and could be prepared in advance. Unlike sports broadcasts, player and game statistics require real-time information transfer so that the visualizations shown are up-to-date and relevant to the current scene for situated sports visualization. Finally, metadata is helpful to provide additional context to the game and enrich the user experience, for example, allowing the expression of emotions by the spectators.

Context

The *context* component ensures that the visualizations appear at the right time and place. The *context* describes the situation in which an entity is in (Grubert et al. 2016). With reference to FW1, we adopted the concept of an object-based and a scene-based *context*. *Object context* primarily refers to details from a specific object, such as visualizations attached to a specific player. *Scene context* is mostly environmental context, such as occurring events and positions of players in the scene. As the user plays a crucial role in many applications, we extended FW1 with the *user's context*.

To ensure that visualizations appear at suitable positions, tracking spectators (*user localization*) in the stadium and players on the field (*scene context*) is essential. Unlike pre-calibrated broadcast systems (Stein et al. 2018), we cannot rely on pre-calibrating the spectators' position in the stadium as even if they are seated in a specific place, the actual position and orientation of their devices vary a lot. Visualizations will not make sense if misaligned from the referent (*canvas*), especially if the actions are reasonably far away. Thus, model-based localization combined with real-time tracking is a suitable solution to calculate the pose of spectators (Baker et al. 2018, Gul et al. 2021). For player tracking, computer vision-based tracking is an option to calculate the positioning of players from cameras installed around the venue.

Other options are wearable technologies, specifically those made for sports analytics (Adesida et al. 2019). These approaches involve using wearable sensors on players, which are then tracked with a receiver to provide highly accurate positioning. Examples would be the sports wearable by KINEXON[1] and ChyronHego[2] that track players' 3D coordinates and measure multiple sports metrics all while having a low latency (e.g. around 20ms for KINEXON).

Upon getting accurate positioning of the user, which is the localization sub-component, we would then need to know when and what to show. This is a combination of both the *scene context* and the *user context* and the *content* component to determine the type of data and visualization to be shown. FW2 uses user localization, referred to as

[1]https://kinexon.com/technology/player-tracking

[2]https://chyronhego.com/products/sports-tracking

"navigator", but lacks the personalization of each user. However, every spectator can have their preference of visualizations; therefore, in our framework, *personalization* is listed under the *user context sub-component*, where spectators can tailor their profile according to what they want to see. This personalized view is related to the user's level of understanding of the sport. A seasoned viewer would probably choose to see more player stats, while a spectator new to the sport would appreciate visualizations explaining what the referee is signaling. Other examples of *personalization* include highlighting specific players on-field, level of detail of event descriptions, penalty scoring predictions, etc.

4.1.2 Applications and Limitations

We designed the conceptual framework to support on-site sports spectating in a known environment. This scenario covers many different sports, from individual sports such as track and field, and swimming to team sports such as basketball, soccer, and many more, including our use case — rugby. Different sports will interpret the framework differently; for example, the *canvas* would be mainly by the pool in a swimming scenario. A valuable past *content* data would be on pacing, possibly the previous world record pace visualized as a line, so spectators could get *scene context* of how fast the current swimmer is compared to the world records. For basketball, the court and the whole indoor stadium could be the *canvas*, and the *content* could be the free-throw ball's path, providing a spatial context of how the ball entered the hoop. All that is needed is to change the model of the physical environment and it could be adaptable to other arena-based sports.

This framework is designed to guide the early conceptual design process of situated visualizations for on-site sports spectating. It, however, does not include the actual design phase detailing how to design a specific visualization, given a particular format. For example, if we have ball possession data for a game, we would already have the *content*, a combination of past and present data. Designers then have to use the *canvas* component to determine where a visualization should be attached, either by a manual identification or an automated plane detection approach. If designers wanted the visualization to be in a spatial environment, it would be considered a world-referenced *canvas* since the visualization is not attached to a single-player or screen space.

4.2 Implementation of Situated Visualization

Our goal is to explore visualization approaches that would benefit sports spectators in the stadium and to evaluate our conceptual framework. We achieved this by implementing two different situated visualization concepts for sports spectating, where *three 'C's* guide our implementation so that we can attach *content* to *canvases* in the proper *context*. We are focusing on mobile devices as it is the most accessible option with the opportunity to be ported to an AR head-mounted display (HMD) when the technology is mature enough for long-term use. We developed two situated visualization options, one based on a standard TV broadcasting style — *Situated Broadcast-style Visualization (SBV)* and the other in the form of an in-situ AR visualization — *Situated Infographics (SI)*. These two options allow us to explore different aspects of familiarity from traditional broadcast and the spatial awareness AR situated visualization can provide.

4.2.1 *Situated Broadcast-styled Visualization*

Due to the limited screen space in *Situated Broadcast-styled Visualization (SBV)*, we mostly display the ***content*** in text or icon form (Figure 4.2, left). The ***canvases*** are display-referenced and mostly presented as rectangles suitable for showing *box-score data and metadata* from the past and present as text-based representations. Some player or game-based statistics could still be visually represented depending on ***context***. Although spectator poses and field players do not need to be tracked for visual alignment (no *identification component* is required), the *context* is still vital for *SBV*, especially *scene context* for appropriate timing of full-screen visualizations. For example, while showing stats during the break, it is possible to utilize the screen space fully and set the whole screen as the *canvas*. The same could not be done if there are still actions on the field and, therefore, would require some intelligent placement of the

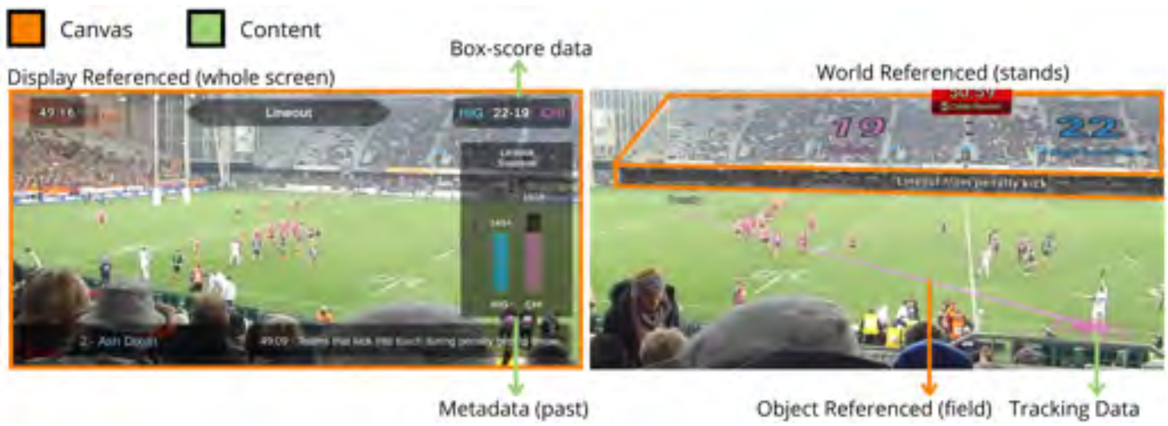


Figure 4.2: An example of *Situated Broadcast-styled Visualization (SBV)* and *Situated Infographics (SI)* alongside references of the *canvas* (in orange) and *content* components in the proposed framework (in green). We are unable to visualize *context* here, but visualization shown are based on user localization and *scene context* of the game. Left: *SBV*. Right: *SI*.

canvases.

4.2.2 *Situated Infographics*

The second visualization approach we developed is an in-situ AR-style visualization called *Situated Infographics (SI)*. *SI* unlock the limitations of restricted screen space as the surrounding environment can be the canvas for visualizations. *SI* work similarly to embedded visualizations, where visual elements are transformed into a 3D visualization and visualized in a 3D space, aligned to their referents. The main goal of *SI* is to visualize complex information coherently and related to the users' environment in real-time. In summary, *SI* is a world-centric visualization method that considers the context around it. It can be seen as a form of the implementation of a WorldBoard concept (Spohrer 1999) where the information is placed in the environment as if it belonged in the real world (Figure 4.2, right).

The ***canvas*** for *SI* is important to ensure that visualizations have somewhere to anchor on in the environment. The canvas could be in the form of a static canvas (*world-referenced presentation*) such as attached to the field, the stands, or in the form of *dynamic canvases* (*object-referenced presentations*) since each player on the field can be a canvas on their own. The ***content*** is similar to the content in *SBV* but contains more elements suitable for graphical representation and that even have an additional dimension of a spatial element. *SI* utilize the *tracking data* to illustrate past and present events on-field better. All information is adaptive to the *canvas* and customized to the viewpoint of the spectator watching it.

The ***context*** is essential for *SI*. Without the *user context of localization*, most visualizations would not work, and any misalignment in the graphics would cause more confusion than assistance. Without context awareness (*scene context*), visualizations may not appear at the right time and distract spectators from their experience. *Dynamic canvases* require anchoring to the players on the field. Personalization and user localization are part of the *user context* component. For an information-rich *SI* implementation for on-site sports spectating, spectators should be able to choose (automatic or interactive) what level of information they want to see. Visualizations should then also adapt to the environment, depending on what is happening nearby and the spectator's perspective.

4.3 Lab Study

Our first goal was to study the general potential of situated visualization within the sports spectating use case. For this purpose, we designed a preliminary user study based on Rugby Union and compared the first design of *SI* to *Traditional Infographics (TI)* on a mobile device. *TI* is the infographics we are used to seeing on paper or on screen. We decided to use *TI* as a baseline as most people are familiar with the concept, and they aim to convey the data in a comprehensible way (Siricharoen 2013).

In a laboratory setting, we applied the *Three C's* to both forms of infographics: traditional and situated. Different parts of the virtual stadium environment served as *canvases* to put dynamic sports *content* into *context*. Since this was a lab simulation, we used past data for the *content*. For this user study, we did not include any user personalization yet; therefore, the *user context* consisted only of *user localization*. The

Figure 4.3: A user study participant using the Lab AR prototype. Also seen is the smartphone (hand-held) to view the *SI* and the other smartphone (on table) displaying *TI*.

user is localized with respect to the lab stadium environment. For *TI*, the *canvas* is mainly display-referenced, whereas for the *SI*, a combined body-world reference model is used.

We were particularly interested in evaluating the user preference for using *TI* or *SI*. We were also interested in exploring whether there are differences in workload between using *TI* vs *SI*, as we would not want to overwhelm the spectators with unnecessary effort. Before undertaking the study, ethical clearance (D19/291) was obtained from the University of Otago's ethics committee and followed the requirements (pre-pandemic).

4.3.1 Design and Apparatus

We designed a within-subject study for comparison to investigate the workload when using *TI* and *SI*. The dependent variable is the workload, and the independent variable is the visualization method with two conditions: 1) *TI* and 2) *SI*. In addition, we collected user preferences feedback on both methods and studied the effect of the visualization techniques on spatial understanding.

For this user study, we used a miniature version of a stadium pitch printed as an A0 poster in a controlled laboratory environment (Figure 4.3). We presented both *TI* and *SI* on mobile phones (Samsung Galaxy S6 and Huawei Mate 20 Pro, respectively) to simulate the participants spectating a game in a stadium. In addition, we used a Microsoft Surface Pro 2017 tablet to capture the participants' responses. This process included marking positions on a photograph of the miniature stadium (captured from the participants' estimated position) and answering questionnaires. We also used historical rugby match data from a sports scoring provider to create a realistic simulation. We used two data sets for different scenarios in a randomized controlled order to avoid learning effects while maintaining the same graphical style for both conditions.

4.3.2 Participants

We recruited participants from the university through advertisements and word of mouth. In total, 30 participants aged between 21 and 38 ($\bar{x} = 26.6$, $\sigma = 4.36$) were recruited for our user study. 23 of the 30 participants were male, and all had a normal or corrected-to-normal vision. Of the 30 participants, 12 stated that they had experienced AR before the user study.

4.3.3 Procedure

After signing a consent form, participants completed a demographic questionnaire requesting information on age, gender, vision impairments, and familiarity with AR. Participants received an introduction to both interfaces, including mentioning the randomized rotation of the field in *TI* to simulate users sitting on the opposite side of the stadium. Upon familiarization with the interfaces, participants were given three tasks to complete, each using either the *TI* or the *SI* condition in a randomized controlled order. We designed the tasks to require participants to understand the infographics and create a moderate workload.

For **Task 1**, we separated the rugby field into five columns and asked the participants to find the area where the most tackles occurred during a rugby game. In the *TI*

Figure 4.4: Lab study: The task here (Task 2) is to mark the position of the orange's team Player 14. Left: *TI*- this version is rotated at a controlled random order compared to what the user sees. Middle: *SI*- infographics directly on the printed field. Right: The printed field with annotations showing wrong (red) and correct (green) responses by the participants.

condition, we showed an infographic with two charts. One chart with colored dots on a 2D picture of a field representing tackles and a second one showing a bar chart with the cumulative tackles for each meter of the field. For *SI*, we visualized the same dots and bar chart except that the dots are on the field in the AR view with the bar chart on the side of the field. We asked participants to mark the field column with the most tackles on the study tablet for both conditions.

For **Task 2**, we visualized the initial position of all players on the field while we asked participants to find the position of a specific player. For *TI*, the initial positions were displayed on a 2D field (Figure 4.4, left) while for *SI* the positions were displayed in AR (Figure 4.4, middle). Participants had to mark the position of a specific player on a photograph of the rugby field on the study tablet. For the *TI* condition, the orientation of the *TI* is randomized to simulate spectators seated in different positions in the stadium. We notified participants in advance before the start of the study to pay attention to the orientation of the infographics. We decided against rotating the printed stadium as it would be obvious that the orientation is being changed and thus decided to go with rotating the *TI* instead.

In **Task 3**, the participants had to find the team with the highest number of votes. For *TI*, we visualized six player-profile pictures (sorted by the team) with the number of votes each received next to each other. For *SI*, we display the players' profiles as banners standing at the side of the field. We asked the participants to select the side of the team with the highest voted player. Participants were asked to complete the subjective workload assessment tool — NASA Task Load Index (TLX) questionnaire (Hart & Staveland 1988) after completing each condition per task. With this we could determine if the workload is any higher than *TI* as a high workload would result in more effort and stress by the users. We then asked additional questions such as preferences and feedback.

4.3.4 Results

We analyzed the workload using the NASA TLX. The results were not normally distributed (tested by Shapiro-Wilk, $p < 0.05$) except for Task 1 *SI* overall condition with a mean of 33.39 ($p = 0.16$). Thus, we analyzed the data with the Wilcoxon Signed Rank test and paired t-test for the Task 1 *SI* overall. The results indicated no significant differences in each category of the TLX questionnaire across the board except for Physical Demand ($p < 0.001$). Therefore, it seems that the participants did not experience a higher workload despite integrating the infographics into their FOV, adding an additional dimension of spatial understanding (Figure 4.5). This is the case for all TLX categories except physical demand, which we assume is likely due to spectators using the mobile AR to look around.

In addition, we found that there seems to be a lack of spatial awareness while using *TI* during tasks that require an understanding of the spatial relationship between the data and the user's environment. For example, for task 2 (demanding the player's position on the field), only 18.7% of the answers were close to the actual position on the field when using *TI*. In contrast to 100% correct answers for the *SI* condition. This is due to the randomized rotation of the infographic simulating the user on different sides of the field. *TI* viewed from the opposite side of the stadium would require a diagonal flip instead of just a horizontal flip for the proper position. For example, the correct position for orange's team player 14 should be in the bottom left corner instead of the bottom right (Figure 4.4). There is no such problem with *SI*, as the spatial relationship is already clearly visible to the user.

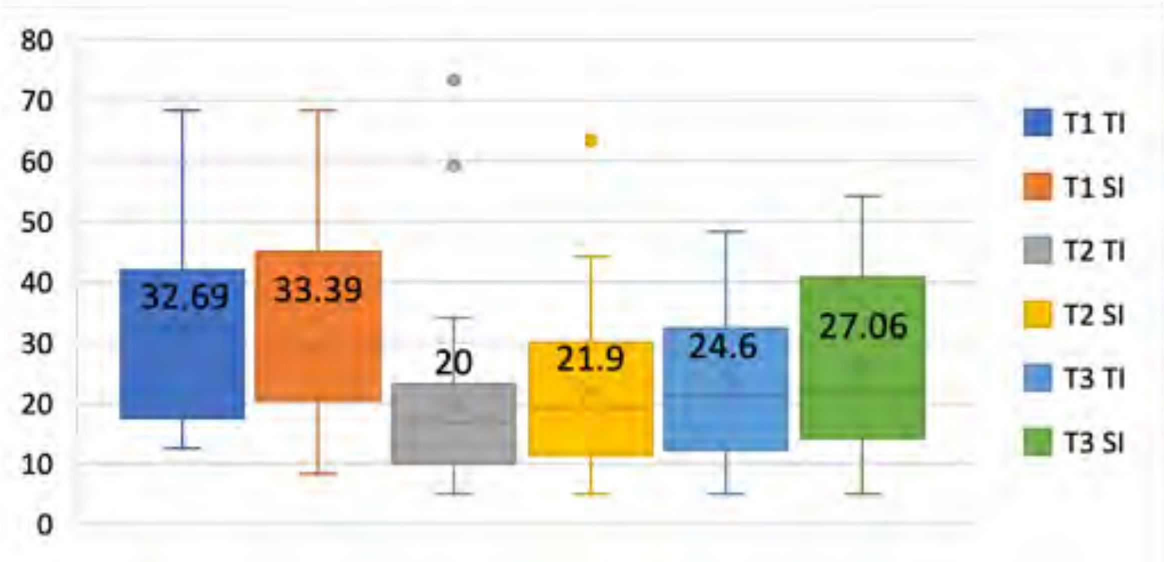


Figure 4.5: Overall TLX scores of all tasks in the lab study. Dot represent outliers and value labels show mean values of the TLX scores.

Finally, we analyzed user preferences. The majority of participants (90%) favor *SI* over *TI* (rated ≥ 5 out of 7, $\bar{x} = 5.74$, $\sigma = 1.28$, where 7 is a preference for *SI*). An analysis of the participants' feedback shows that 26 out of the 30 participants (86.6%) provided positive feedback towards the idea and concept. The words "useful", "fun", and "interesting" appeared rather frequently in the feedback (Figure 4.6). Fifteen participants mentioned *SI* as useful or helpful to them, some stating that it is useful for newcomers to the game and hard-core fans who wanted statistics quickly. Participants also mentioned that the mixing of the visualizations with the environment is pleasing. This indicates that users are pleased with the concept of *SI*.

Regarding some drawbacks mentioned, participants wanted a better-polished front-end implementation, bigger fonts, and simpler infographics. A participant recommended using symbols instead of numbers. For instance, the number of votes obtained for each player could be replaced by star ratings. Some participants also reported that some infographics are easier to see under *TI* conditions.

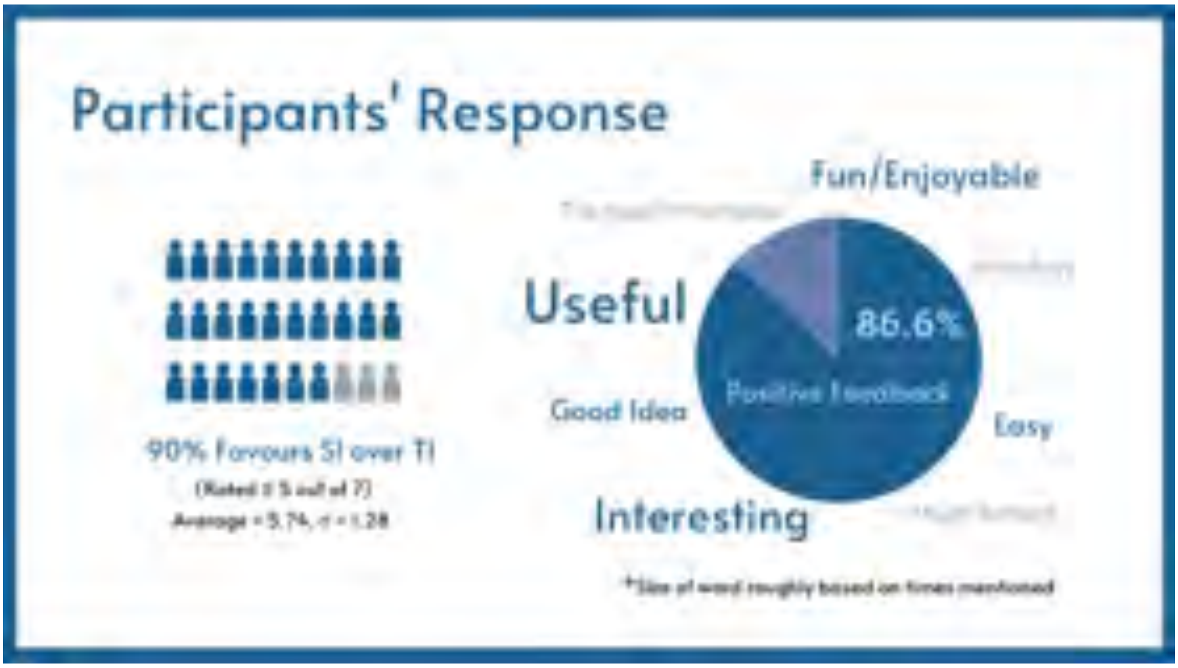


Figure 4.6: A word cloud infographic of participants' feedback in regards to the user study.

4.3.5 Discussion

The preliminary lab study indicated that *SI* could help with spatial understanding without significantly increasing workload compared to *TI*. Considering that *SI* is used in a more complex environment, it brings together more benefits, including monitoring what is happening on the field in real-time and providing better spatial understanding while still keeping the cognitive load similar to *TI*. Therefore, these results suggest that, although they have a similar cognitive load, users are getting more out of *SI*. We also found that the participants favored *SI* in their preferences and emphasized aesthetics and easy-to-understand graphics.

4.4 Formative On-site Study

Our lab study indicated an interest in AR situated visualization sports spectating and that there could be benefits compared to more traditional ways of presenting information on-site. Thus, we ran an on-site user study in a stadium (Figure 4.7, left)

Figure 4.7: Left: Explaining the prototype to the participant in the on-site user study. Right: Figure showing the two visualization approach we developed - *SBV* and *SI*.

to evaluate situated visualizations on-site (Figure 4.7, middle and right) as well as user preferences towards different spectating conditions. The experiment was divided into two sessions: an on-site study in the stadium and an off-site session that simulates a spectator using traditional television broadcasts. Following the pilot study, we again focused on Rugby Union for this study. This study has received ethical approval (D20/254) from the University of Otago's ethics committee and followed health and safety precautions (pandemic).

4.4.1 Design and Apparatus

Similarly to the lab study described above, all *three C-components* were applied: We localized the users within the stadium environment and visualized match-related *content* (present content). The presentation was driven by a full-fledged combination of world and object references for the *SI* condition. To make the *SBV* condition ecologically valid and fair comparison, display-referenced data has been combined with *scene context*, providing visualizations at the appropriate time. *Canvas* identification was pre-modeled, i.e. a manual approach was used.

The on-site user study is held at the Forsyth Barr Stadium. We decided not to run the experiment during a live match to avoid bias from an environment such as sudden disruption from the crowd and events that are very difficult to control and also due to restrictions concerning the COVID-19 pandemic. For the on-site visualization, we used an iPhone XR for all parts of the study. We captured 360° video footage during a live rugby game for the indirect AR prototype. For the lab session, we showed the broadcast footage of the same game on a 32-inch monitor with participants seated on a sofa approximately 2m away.

4.4.2 Assumptions

Based on the design considerations, we postulate the following assumptions for our on-site study which is related to **RQ3**.

A1. We do not expect a difference in presence and user preference between Indirect AR and AR for our on-site spectators.

A2. We expect a similar workload for SBV and SI.

A3. User experience for SBV and SI will be above average.

A4. Game understanding will be supported by SBV and SI in a similar way to watching a broadcast and will increase compared to watching a game on-site without any support.

4.4.3 Participants

We recruited participants from the university through advertisements and word of mouth. The selection criteria are that participants must have been to at least one game in the stadium before. In total, 16 participants (11 Males, 5 Females) aged between 19 and 36 ($\bar{x} = 24.9$, $\sigma = 5.12$) participated in our study. Among the participants, eight claimed that they had prior experience with AR, four claimed that they had no experience with AR but knew about it, and another four had no experience. Four participants who majored in sports-related degrees have many sports spectating experiences and are deemed expert users.

4.4.4 Procedure

Participants were first invited to read the information sheet and then asked to fill in the demographic questionnaire, the COVID-19 declaration and the consent form. We designed the on-site study to consist of two parts.

Part 1: Comparison of AR and indirect AR

The first part compares actual AR and a video indirect AR, which is a prerecorded 360° video of the empty field in the stadium. Both conditions feature the situated visualizations we used in the pilot study to reduce confounding variables. We gave participants time to explore their environment and visualization for at least half a minute until they were satisfied. They were then presented with a questionnaire consisting of some relevant questions selected from the *igroup presence questionnaire (IPQ) overview* (n.d.) (refer Appendix A) and asked if the conditions affected what they saw. This will assess our assumption A1.

Part 2: *SBV* and *SI* comparison

Participants then entered the second part of the study with two different situated visualization conditions, both presented in the indirect AR prototype. For this purpose, we used footage from a rugby game that we captured on a 360° camera (Insta One X). The first visualization condition is *SBV*, a broadcast-like overlay visualization similar to what spectators experience on television with scores, timers, and visualizations overlaying on the video image. The second visualization condition is an updated version of *SI*, which places the visualization into the environment. This means that visuals could appear anywhere, such as in the field and the stands (see Figure 4.7, right). The participants were given two tasks for each condition. Each task consists of a spatial component and a game stat or game understanding component. We asked participants to answer questions related to each task such as *"Which of the two condition they find more appealing?* "after the video clip ended, which lasted around a minute and a half. The main purpose of the task is to have the user focus on a task so that the NASA TLX questionnaire (Hart & Staveland 1988) can be applied later to assess assumption A2.

Task 1: Determining where a specific field action (a penalty) occurred and what caused it.

Task 2: Locate the initial position of a certain player and identify the team with the highest ball possession.

The video clip presents all the events and related information in both conditions. The clips and order of conditions were all in a randomized order. The participants did not view the same clip for different conditions. We also gave participants a hidden task regarding the line-out success rate and players' roles that also appeared in the visualization but were not written in the task list. Participants filled out a NASA TLX questionnaire after each task. Then they were presented with both conditions again for the last time. After each condition, they completed a User Experience Questionnaire (UEQ) (Schrepp 2019) to rate the visualization method without focusing on any task. The questionnaire continued to ask them about their preferences and experience which assesses assumption A3.

Part 3: Comparison of broadcast footage viewing

The last part of the study involved viewing a short snippet of a broadcast game on a different day in the lab after the on-site study for assumption A4. Participants received a 5-minute broadcast video of the game they saw in the stadium through the video indirect AR in a simulated home viewing environment. They were then asked to complete a questionnaire on all viewing conditions they had experienced before, from on-site viewing without assistance, SBV, SI, and finally, the TV broadcast. We asked additional questions based on their response to get more feedback.

4.4.5 Results

In the following, we present and discuss the results for each of the three parts of the study.

Results Part 1: AR and Indirect AR comparison

For comparing the AR condition with the indirect AR condition, we used a subset of the Igroup Presence Questionnaire (IPQ) test. Only two questions (Q6 - I felt present in the augmented environment. and Q7 - I was not aware of my real environment.)

passed the Shapiro-Wilk test of normal distribution ($p > 0.05$), which then underwent the paired t-test ($p = 0.61$ and $p = 0.84$). We used a Wilcoxon signed-rank test for the remaining questions ($p = 0.85, p = 0.16, p = 0.80, p = 0.49, p = 0.40$). For the seven questions, we found no significant differences. Similar findings were captured when participants were asked about their preferences. Of the 16 participants, 9 preferred the AR condition and 7 preferred the indirect AR condition. Also, when asked if it matters on which condition they view the visualizations, participants rated that it does matter slightly ($\bar{x} = 4.06, \sigma = 1.91$ on a 7-point Likert scale in which higher means it matters more). Some participants said that the AR condition felt more realistic (P1, P4) and sharper (P2, P3, P8) than the indirect AR, but some stated that the indirect AR is more pleasing (P5, P10, P14) to look at. These results confirm our assumption A1 in that we could not measure a significant difference between AR and Indirect AR for our sample spectators.

Results Part 2: *SBV* and *SI* comparison

In the second part of the experiment, we compared the two situated visualization approaches, *SBV* and *SI*. We used the TLX test and collected 15 usable data entries for the TLX scores (one participant who rated almost a minimum score for every category was considered an outlier and was discarded). The overall TLX scores passed the normal distribution test using the Shapiro-Wilk test ($p = 0.58, p = 0.93, p = 0.25, p = 0.72$) and then underwent a paired t-test. Overall TLX results show no significant differences between the two visualization methods (Task 1 $p = 0.09$, Task 2 $p = 0.44$). This confirms our assumption A2, indicating a similar workload for both conditions for our samples. However, when separated into normal users and expert users, there is a significant difference in the overall mental demand and effort of normal users for Task 1 ($p = 0.03, p = 0.03$ and $p = 0.04$) with a lower average TLX score for *SI* (boxplot Figure 4.8). It is important to note that the workload scores for both conditions are lower than the 50% of previously observed Computer Activities (54.0) and lower than the 75% of previously observed Visual Search tasks (51.06) (Grier 2015).

The UEQ scored well for both conditions with positive scales on all aspects. We

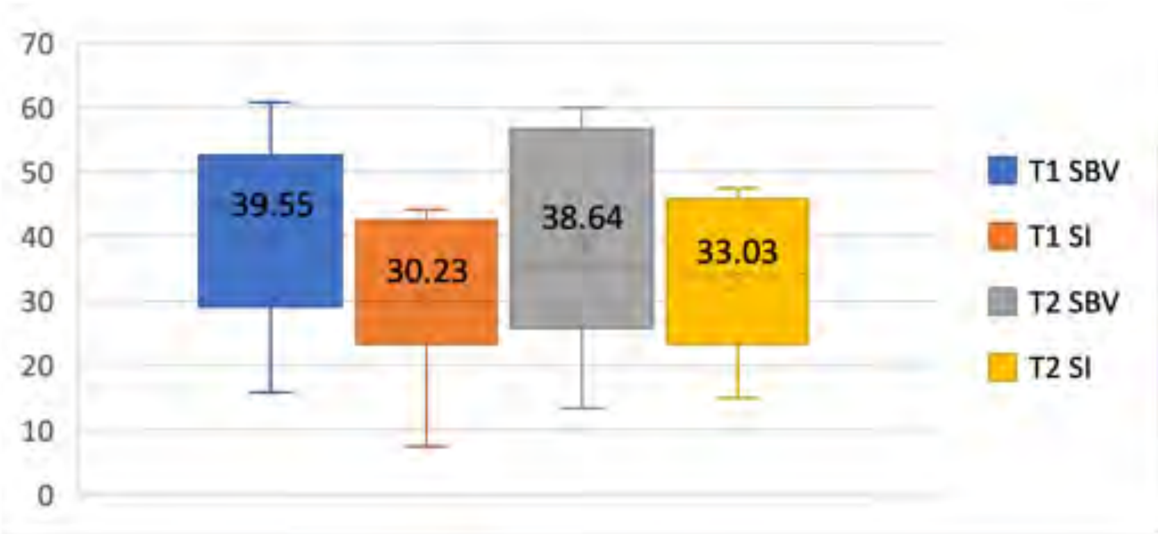


Figure 4.8: Overall TLX scores for the 2 tasks (T1, T2) in the on-site study. Value labels show mean values.

could see that all average *SI* scores are higher than those in the *SBV* condition. However, we could not measure any significant differences between the conditions in the six categories with a two-sample t-test assuming unequal variances ($p = 0.53, p = 0.37, p = 0.96, p = 0.90, p = 0.06$ and $p = 0.10$). We also compared our UEQ scores for both visualizations approaches with the benchmark provided by the UEQ Data Analytics Tool (version 4) (Figure 4.9). For SI, most scores are above average and excellent among the benchmark, while the SBV ratings are mainly in the below-average range. The benchmark is a growing collection of all UEQ evaluations that are shared by contributors (Schrepp et al. 2017). According to the UEQ handbook (version 8) (Schrepp 2019), there are 452 products in the benchmark with 20190 total participants. The results indicate that we can confirm hypothesis A3 only partly for *SI*, and we need to reject it for *SBV*.

When asked about preferences for on-site spectating, nine participants preferred *SI* and seven preferred *SBV*. Among the reasons for choosing SI are that it is not too cluttered (P7, P8, P12, P13), easy to understand (P4, P10), well-integrated/seamless/realistic (P4, P7, P11), and it is nice to have information on the field (P6, P12). Regarding *SBV*, participants mentioned that it does not distract from the game as the visuals are not on the field (P1, P14, P16), it is closer to real-life (P1, P2), and is clearer/simpler

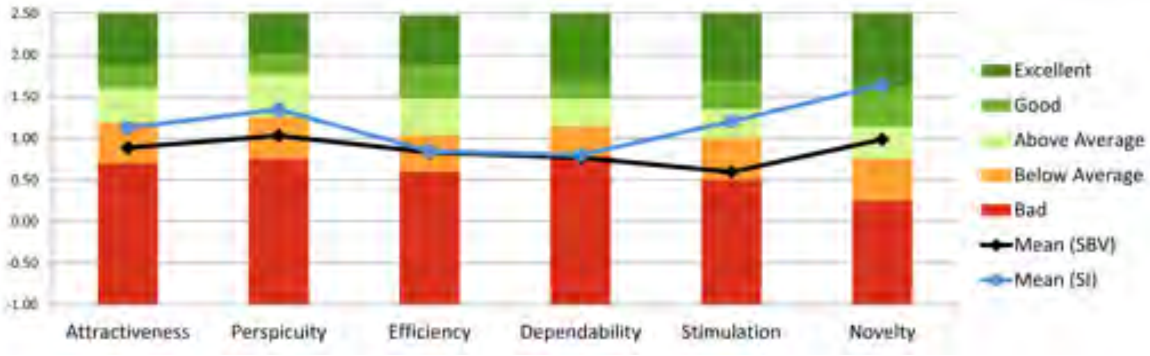


Figure 4.9: UEQ benchmark score (version 8) comparison of both visualization approaches, situated broadcast-style visualization (SBV, black graph) and *SI* (SI, blue graph).

Figure 4.10: Ranking of sports spectating method among aspects.

(P3, P14, P15). There are some contradicting statements among the participants as some said that the *SBV* is more minimalistic than *SI*, but others said otherwise.

Results Part 3: Broadcast Footage Viewing Comparison

An overall comparison of all methods, including off-site spectating, indicated that *SI* is the first ranked condition for game understanding: higher is better ($\bar{x} = 1.8$, Figure 4.10). The Friedman test rejects the null hypothesis showing that differences are found in overall game understanding ($p < 0.001$) and overall team fanaticism ($p < 0.001$). Overall team fanaticism here refers to the feeling of attachment to the participants' supporting team, which provides a good indicator on how immersed they are into the game. A further post-hoc Nemenyi test shows that in both overall game understanding and team fanaticism, there are significant differences between unassisted on-site viewing and the rest of the conditions (Game understanding $p = 0.04, p = 0.01, p = 0.03$, team fanaticism $p = 0.01, p = 0.01, p = 0.02$). This is in line with almost everyone agreeing that on-site viewing without assistance was the worst in terms of game understanding (12/15 participants, Binomial Test with CI $p = 0$) but the best for overall team fanaticism (13/15 participants). These results confirm our assumption A4, indicating that the situated visualization approaches increase game understanding compared to the absence of assistance.

4.4.6 Discussion

Considering that we did not find significant differences in IPQ and the ratings varied between the actual AR and the indirect AR, it seems acceptable that indirect AR could be used to represent the actual AR in our setting, although more research is needed. When comparing the two situated visualization methods, we found improved game understanding and good scores in UEQ, with *SI* leading *SBV*. Situated visualizations seem to improve the overall experience, except for team fanaticism, which is only enjoyable without help from visualizations.

However, we found that different users prefer different visualization methods. When it comes to *SBV* and *SI*, the views of the participants were mixed. Half of the par-

Figure 4.11: Hybrid visualization prototype based on the user feedback from the on-site study.

ticipants (P4, P7, P9, P10, P11, P13, P14, P15) felt that a personalized view was necessary due to too little or too much information.

Some participants (P5, P13) suggested a hybrid visualization approach. Some mentioned that it would greatly reduce the physical demand, as users, for example, do not need to keep moving to the center to see the timer (P13).

Most of the participants focused exclusively on the given task. When asked additional questions, the success rate was relatively poor. Some participants mentioned that they noticed other visualization elements but did not pay much attention as it was not part of the current task. One participant (P5) said that the line-out visualization for T1 was too quick and clashed with another ongoing event, causing distractions. This also indicates the importance of having interactivity and personalized views.

4.5 Conclusion

In this chapter, we explored situated visualizations in the context of on-site live sports spectating in a stadium environment using rugby games as examples. We presented a new conceptual framework for on-site sports spectating that guided the implementation of our prototypes and the design of our studies. Although White's original framework for situated visualization (White & Feiner 2009*b*) served as a good starting point, the findings of our study confirm the need for a comprehensive framework considering additional aspects needed for the visualization of sports data on-site. Our *3'C's Situated Visualization for on-site Sports Spectating Framework* showed its usefulness when comparing dynamic AR *content* on different canvases within the stadium *scene context* where the user is in complete view control with the combined approach of a display-referenced *canvas* for *SBV*.

We conducted two user studies that support the general feasibility of our approach and found a preference for situated visualizations compared to just watching on-site games. Spectators are getting more information from *SI* compared to *TI* while maintaining a similar cognitive load. We also show that situated visualizations generally assisted in overall game understanding while maintaining the game satisfaction higher than just watching the TV broadcast. An interesting observation is that there is no right or wrong approach when choosing between *SBV* and *SI* as depending on individuals, both are equally favorable. However, we implemented a hybrid prototype that combines *SBV* and *SI* (Figure 4.11) to get the best of both worlds as a starting point for future work.

Chapter 5

Technical Factors Affecting the User Experience

Contents

In this thesis, we have explored the components of the *ARSpectator* system and methods to design situated visualizations for on-site sports spectators. There are two main components of the simplified overview (Figure 1.3) in Chapter 1 that were not discussed — the content sources and the user tracking. As we are researching visualization and user experiences, we were interested in investigating whether these two components could impact user experience.

5.1 Motivation

Despite advances in hardware and software development, most AR applications still focus on augmentations in small workspaces (Klein & Murray 2007, Bach, Sicat, Beyer, Cordeil & Pfister 2017, Akçayır et al. 2016), and there is limited research on AR for large-scale environments. Research often separates technical factors, such as tracking accuracy or success, from usability of specific aspects of an AR application (Schall et al. 2013, Gul et al. 2021). However, it remains unclear what influence well-known technical factors have on the overall user experience. Therefore, in this chapter, we seek to address this gap by analyzing the technical factors that influence user experience in large-scale AR environments, specifically in a stadium for a sports spectating context.

The *ARSpectator* system allows on-site spectators to retrieve game-related information from our content source, a hosted server. The server provides real-time game statistics obtained from the sports data provider and player tracking information generated by a vision-based player tracking system (Zollmann et al. 2019). This general architecture is typical when providing similar information to a live TV audience. It introduces different technical factors that could affect the AR experience. From preliminary testings, we noticed that these factors are latency, registration accuracy, and jitter, which originate from both the client device used for the AR experience and the dynamic content source. The factors affecting the client device are mainly caused by camera registration and tracking inaccuracies (Zollmann et al. 2019). In contrast, the factors of the content server are similar but would only affect dynamic content visualizations.

Although it is a common assumption that technical factors influence the user experience in AR, so far, it remains unclear to what extent they have influence and whether there are factors that have a greater or lesser impact. Based on this gap, we identify an essential question that, to our best knowledge, has not been explored or answered so far: How do technical factors, such as system limitations or errors, affect the user experience of a large-scale environment AR system, particularly if dynamic content is presented? With that question in mind, we expand **RQ4** *(How do technical factors in a large-scale environment AR affect the user experience of the sports spectators on-site?)*

with the following research questions.

RQ4-1. Among the selected technical factors, are there any factors that are more disruptive than others to the user experience in a large-scale environment AR application with dynamic content?

RQ4-2. For each of the technical factors, what are the noticeable levels and disruptive levels of those factors where spectators might start to be negatively affected in terms of user experience?

RQ4-3. If we were to combine the noticeable levels of the different factors that affect the user experience, would the effects accumulate?

We formulate these research questions as there is a lack of research in measuring acceptable and disruptive values for the technical factors, especially registration accuracy and jitter in a large-scale environment. Previous research looking at similar factors measures the effects of mental load and task performance (Henze & Boll 2010) or the ability to perform specific tasks (Rogers et al. 2017). This differs in a sports spectating context as spectators mostly spectate a sport for entertainment, interest in a specific sport, a change of mood, to support a team or someone they know (Apostolou & Lambrianou 2017). This finding shows that there is no motivation to finish specific tasks in a set amount of time or with a certain accuracy. Therefore, we decided to shift the focus to user experience in our study instead of performance metrics.

5.2 Background

5.2.1 Technical Evaluations in AR

As identified by Kim et al. (2018), tracking and localization methods have been considered the most popular topics of AR research for more than two decades. A considerable body of computer vision-based research looks at different algorithms for SLAM, an algorithm that allows a machine to map out its environment while trying to localize itself (Reitmayr et al. 2010, Liu et al. 2018). We implemented SLAM-based methods for our stadium environment (Baker et al. 2020, Gul et al. 2021). While tracking and localization are essential technical factors to consider for user experience, these papers usually feature technical accuracy results that do not translate to user experience measures. From this technical research, we derive registration accuracy and jitter as technical factors to be used in our user study.

Latency is an additional factor that can be influenced by an AR application's tracking and localization methods. Computationally expensive methods will add to latency as they require a longer processing time. In addition, latency can result from AR HMD latency (Jacobs et al. 1997, Lincoln et al. 2016) or from AR simulation (Lee et al. 2010). Latency reduction also has been investigated in recent years (Chen et al. 2018, Lincoln et al. 2016). Braud et al. (2017) examined how networking technologies could be improved to support the massive offloading of mobile AR applications that introduce latency to digital content visualizations. However, similar to research on tracking and localization, most of this work on optimizing latency is technically focused and does not involve users. Therefore, how much latency affects the user experience is not yet clear.

5.2.2 User Experience Evaluations in AR

A survey by Dey et al. (2018) shows that user evaluations are done targeting interaction and perception. However, many of them focus on performance-based evaluations, such as using the NASA TLX questionnaire to determine mental load (Henze & Boll 2010). The user experience usually goes beyond this and includes the usability, pragmatic, and hedonic aspects of a system (Ritsos et al. 2011). However, usability metrics and acceptance questionnaires do not assess experiential and emotional aspects (Olsson 2013). Some research was carried out in terms of summarizing user experience evaluation methods (Vermeeren et al. 2010), but, similar to what Olsson (2013) reported, none of this research is actually used in the context of an AR use case.

Some studies have a few questions asking if participants find the process enjoyable (Henrysson et al. 2007). A user study evaluated technical factors such as latency, framerate, resolution, and jitter to find effects on presence (Louis et al. 2019). Other work that have questionnaires on user satisfaction are more often usability ratings rather than user experience (Fonseca et al. 2014). Iwata et al. (2011) use intrinsic motivation (Ryan 1982) to measure learning motivation, which also includes some form of user experience in the category of interest/enjoyment subscale. They also conducted a free-writing style survey to get better insight from the participants. We think that the above studies are not sufficient for any form of actual user experience evaluation as the primary motivation of their research was not on user experience but instead just collecting some user experience-related feedback from participants.

5.2.3 AR Design Guidelines

In previous work, various authors have studied AR design guidelines that focus on providing a better user experience (Ortman & Swedlund 2012, Ganapathy 2013, Vi et al. 2019). These guidelines, however, emphasize design factors rather than technical factors. While some AR design methods are proposed, they lack formal user evaluations (Barba et al. 2010, Vi et al. 2019). Dünser et al. (2007) applied the human-computer interaction design principles to the design of AR systems and identified a set of design principles that could also be used for AR systems. The authors mentioned tracking stability and registration error issues as potentially contributing factors to user frustration but did not conduct any user studies on these factors.

5.2.4 Mixed Reality Solutions for AR Evaluations

It is difficult to conduct reliable AR evaluations due to many technical aspects, such as constant changes in lighting or placement of items in the real world (Renner & Pfeiffer 2017). Simulating AR using VR has been explored for quite some time. Due to complete control of a virtual environment, different factors can be isolated and manipulated individually, allowing evaluation formats that are not possible in an AR scenario (Ragan et al. 2009, Renner & Pfeiffer 2017, Hettig et al. 2018). Mixed reality setups can also provide a simulated AR experience that is not yet technologically possible, such as a wide field-of-view AR (Ren et al. 2016). In some cases, a mixed reality approach, such as a CAVE system, is more accessible than an AR approach for evaluation purposes (Tiefenbacher et al. 2014) because the researcher can observe what the participants see. Despite this, there are challenges such as incorrect depth perception and the lack of tactile feedback from the real environment with mixed reality solutions to evaluate AR (Ragan et al. 2009). However, in our use case, where we want to evaluate and manipulate isolated technical factors in a large environment, a mixed reality solution is still the best way. We implemented an indirect AR approach (Wither et al. 2011) for our user study.

5.3 Technical Factors Affecting the AR User Experience in Sports Spectating

Olsson & Salo (2011) conducted a large-scale online survey that involved more than two thousand participants in AR mobile applications. In their survey, users identified technical and functional problems as the main weaknesses of AR applications. In particular, as an example, they often mentioned imprecision in localizing as a technical problem. This is related to our target registration accuracy and jitter factors. Additionally, latency is one of the main factors in AR applications (Holloway 1997) and is also associated with software deficiencies (for tracking) and hardware limitations (networking infrastructure and device limitations).

Latency, registration accuracy, and jitter are common technical factors in an AR sports spectating scenario. In our research, we focus mainly on the two sources of these factors — the mobile client device (device camera) and the player tracking system (Figure 5.1). We did not include the sports data provider as they would also have their technical factors based on their system, which is well outside our control. More potential factors could affect the AR user experience, but we want to focus on the factors caused by technical limitations. Therefore, this thesis does not consider the design or human factors such as visualization style, information cluttering, depth perceptions, ergonomics, etc.

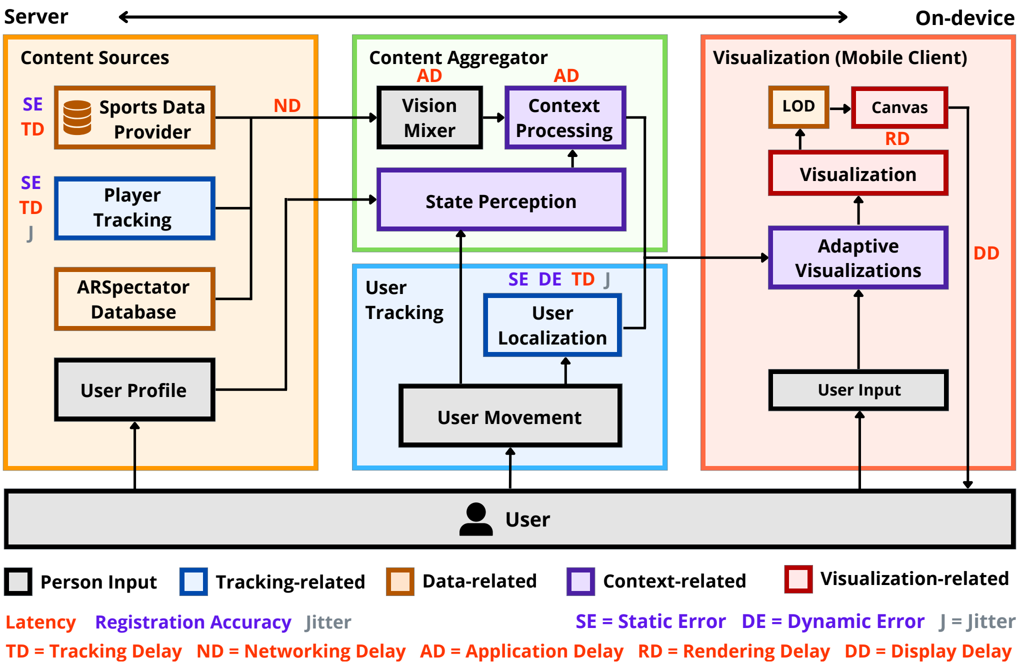


Figure 5.1: The various parts of the *ARSpectator* overview where the three technical factors that we are investigating exist. The three technical factors are latency, registration accuracy and jitter.

5.3.1 Latency

Sielhorst et al. (2007) described latency in VST AR systems as the difference in time from augmented visualization to reality. Latency is more a perception by the end-user, which is equal to system delay (the actual end-to-end delay of the entire system) if no corrections are made to the system (Jerald 2004). Jerald (2004) also combines the research of Mine (1993) and Olano et al. (1995) to describe the sources of system delay in a VR system, which is mostly similar to VST AR systems. The five sources of system delay in temporal order are as follows, written with respect to our *ARSpectator* context (Figure 5.1):

1. *Tracking delay* - This is the delay from the tracking system. It is present in all areas that involve tracking, from the sports data provider with their in-house tracking system, our vision-based player tracking system, and lastly, our user localization component where we track users' pose and position in the stadium.

2. *Networking delay* - This is the delay in communicating data over the network. For

networking delay, it is found mainly in the content source, most prominently from the point at which the content source is transferred to the content aggregator.

3. *Application delay* - The *application delay* is specific to the task performed and not coupled with the other delays. In our case, this is mostly happening in the content aggregator, where the vision mixer might have some delays in annotation or context processing trying to figure out what users would like to see based on the user state and incoming information.

4. *Rendering delay* - This is the delay in rendering the graphics via the graphics component in the hardware. It depends on the scene's complexity and is variable at different times. In *ARSpectator*, it would mean the delay in producing the scene with all the visualizations. Rendering display will be higher with a VST AR application as the whole environment would need to be rendered as opposed to an OST AR setup.

5. *Display delay* - Last, we got our *display delay*, which is present in the final step where the hardware itself presents the visualizations to the user in terms of turning on and off the correct pixels. This should be around 16.7 milliseconds or less in our use case, as most display hardware now displays at least 60Hz, with some VR HMD reaching more than 100Hz.

While we identified many sources of latency above, in this chapter, we are interested in mainly investigating the relative latency of player tracking, which is the latency between two streams of data that causes misregistration, also known as lag (Jacobs et al. 1997, Wloka 1995). The relative latency would include the *tracking delay* to get the players' position on the field and *networking delay* to get all the information to the mobile client. The artificial latency (Lee et al. 2010) we are trying to simulate is between the delay of the tracking data stream and the actual scene captured on camera. This latency will affect dynamic content on the field, especially visualizations that follow a player, such as an arrow or a highlight (Figure 5.2, left).

We decided to only focus on the relative latency of the player tracking because we cannot fully simulate the OST experience nor the VST AR experience with an indirect

Figure 5.2: An illustration of technical factors evaluated: latency, registration accuracy, and jitter. Visualizations are shown in blue (static) and orange (dynamic) where the arrow is targeting the person in yellow.

AR setup. We cannot eliminate the display latency from camera capture to screen, which is around 100 to 150 ms (Chen et al. 2018). Hence, if we were to simulate this camera latency onto an indirect AR, it would introduce many other confounding variables related to compound latency. Besides that, even if we were to consider the latency of the VST AR approach, there has been research done on the end-to-end system latency of VR systems that states anything more than 50ms is perceptible, with anything more than 150ms damages the illusion of presence (Brooks 1999). To narrow down the scope of our user study, we also ignored the *application delay*, *rendering delay*, and *display delay* and only focused on the *tracking delay* and *networking delay*, which are typically similar throughout and could be more reliably simulated in our user study prototype.

5.3.2 Registration Accuracy

Registration accuracy in our scenario aims to mimic the inconsistencies in tracking and initialization. In real applications, this could be due to a misalignment in the initialization/localization phase. However, this is often present, especially in image-based registration. Kruijff et al. (2010) did mention that the human brain can handle inconsistencies in such registrations, and an approximate registration might be sufficient. Therefore, we would like to determine how much inconsistency would be acceptable.

There are two types of registration errors, namely *static errors* and *dynamic errors*

(Azuma & Bishop 1994). *Static errors* are registration errors that occur even when the user is stationary and not moving. For our use case, this translates into registration errors from the content source — the sports data provider and player tracking, and the user localization (device camera). On the contrary, *dynamic errors* occur if the user moves their head, usually caused by the temporal mismatch between the user's head movement and the display. This error would only occur in the user localization part, where the display lags behind the users' movement. We are only interested in the *static errors* for our case. Similar to latency, we are not evaluating the technical factors caused by head movements but rather the technical factors related to position tracking.

The registration accuracy is stimulated by a consistent offset of the visualization from the original location, reflecting an error in placement or alignment. The accuracy of the client device camera registration affects every visualization in the environment. In contrast, the accuracy of the player tracking registration only affects the dynamic spatial visualization showing, for instance, the positioning of the players (Figure 5.2, middle).

5.3.3 Jitter

Jitter is a slight irregular movement of visualizations due to precision errors in tracking and anchoring of visualizations. We could observe two types of jitter in XR systems, the first being *spatial jitter* and the other being *temporal jitter*, which is a form of invariant latency (Stauffert et al. 2018). *Spatial jitter* is most commonly caused by noise in the device signal and the hands of the users themselves (Teather et al. 2009). In our *ARSpectator* use case, this translates into the noise of the player tracking software, the localization of the user, and the tremor of the mobile client device, be it a mobile phone or HMD. *Temporal jitter* on the other hand is jitter due to a change in latency that is not constant, causing visualizations sometimes to be delayed and other times not. Although we acknowledge that *temporal jitter* would most likely be present in our sports spectating scenario, since it is more towards the research area of networking, we decided not to investigate *temporal jitter* in our research. Hence, the jitter we simulate in this chapter are all *spatial jitter*.

We could reduce *spatial jitter* by smoothing out the visualizations. However, that would induce more latency (Louis et al. 2019). As we just mentioned, *spatial jitter* in our system is caused mainly by noise and tremors, which also originate from the device camera and the player tracking system. While there should be minimal tremors on the player tracking system due to its stability in mounting, there would be tremors from the users' movement. Despite that, we cannot simulate jitter originating from the users' device as the users themselves would be holding or wearing it. Therefore, while we are aware the users' arm or head would introduce *spatial jitter*, we only simulate *spatial jitter* caused by noise in the signal of player tracking or user localization. This noise-induced *spatial jitter* from the device camera would affect all visualizations (static and dynamic), while *spatial jitter* from player tracking would only affect dynamic visualizations (Figure 5.2, right).

5.4 Evaluating AR UX Factors in Sports Spectating

To explore what we have discussed in the previous section, we designed a user study using mobile indirect AR to evaluate a selection of technical factors in AR with respect to user experience. We decided to use the mobile prototype because the majority of the general public would be familiar with the interactions of a mobile phone. To avoid confounding factors during live sports games, such as differences in the events, we decided to replicate the AR stadium experience in a laboratory setting; hence, we chose the mobile indirect AR prototype. As mentioned in Chapter 3, indirect AR could be used to support off-site development and evaluation, supported by the indirect AR and AR comparison we did in Chapter 4. The indirect AR prototype provides a more accurate content alignment, effectively controlling latency, registration accuracy, and jitter parameters. This study has received ethical approval (D21/162) from the Ethics Committee of the University of Otago and followed the health and safety precautions (pandemic).

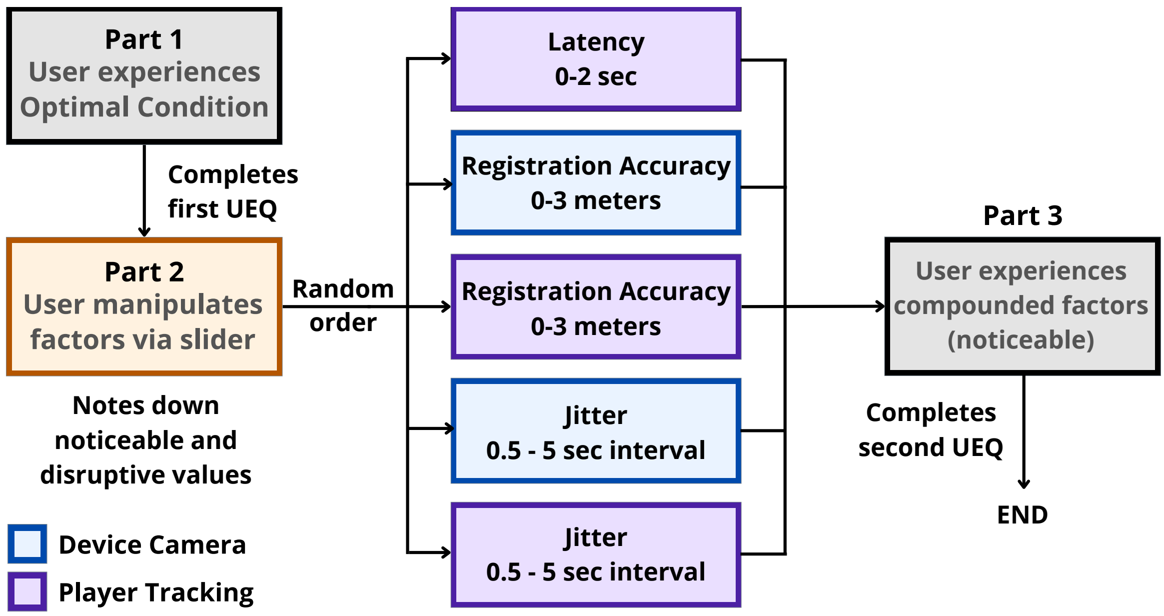


Figure 5.3: The flow of the user study investigating the noticeable value and disruptive value of the three technical factors — latency, registration accuracy and jitter. In all cases, except latency, we look into two sources; the device camera and the player tracking system.

5.4.1 Study Design

In order to evaluate if the specific type of latency, registration accuracy, and jitter described in the technical factor section affect user experience, we designed a within-subject study in which participants have the freedom to manipulate these factors. Apart from determining which factor has the greatest negative impact on user experience, we were also interested in seeing the noticeable and disruptive levels (values) for each technical AR factor. We define the noticeable level as the level where participants noticed an impact on their user experience, while the disruptive level is when participants feel that it is almost to a point where application is unusable. There are two dependent variables for this study, 1) user experience and 2) value in which a factor is noticeable and disruptive. The independent variables would be the three technical factors.

There are three parts to the user study, as illustrated in Figure 5.3. The first part is to experience the prototype in an optimal condition. The second part of the study is the phase where users could manipulate AR factors to obtain noticeable and disruptive values. The third part is similar to the first part; however, it will now be an experience of compounded factors from the device camera (blue factors in Figure 5.3), except for latency, where we only evaluated the player tracking. All three parts involved participants experiencing a looping one-minute indirect AR experience using a 360° video of an actual rugby game with visualizations overlaid onto it. The visualizations show events and details related to the game, such as the current score, game timer, and line-out statistics, for example. The first and last phase use the same 360° video for comparison, while a different clip is standardized for all factors during the manipulation phase. Participants do not need to do any specific task other than to observe the game as that closely mimics what happens in real life. Participants fill out the User Experience Questionnaire (UEQ) (Schrepp 2019) after the first and last part for comparison purposes.

For part two, which is the manipulation of factors, the technical factor and a brief instruction statement are stated at the top of the screen. At the same time, there is a slider and a button at the bottom of the screen (Figure 5.4). The slider controls the degree of linear manipulation of the specific factor, while the button next to it is for the participants to press when there is a noticeable effect on the user experience. This approach has been used before where participants can control specific parameters and indicate what seems perceptually most appropriate to them (Sutton et al. 2022). The slider would then reset, and the wording on the button would change to "Disruptive", where participants would manipulate it again until it represents a disruptive user experience for them. When selecting the disruptive value similarly to how it was done for noticeable values, a rating panel appears, allowing participants to rate how much the factor negatively impacted their user experience from 1 to 7. The prototype is then reset to the initial stage with a different randomized technical factor. We repeat this throughout the second part of the user study. All values selected by the participants, including the ratings, are automatically output into a text file for results analysis later on.

Part three is an identical clone of part one, except that the noticeable value of all the device camera factors and the latency of the player tracking system is incorporated

Figure 5.4: User study session where a participant controls a slider to manipulate the technical factor. The goal is to set the slider until they obtain the noticeable and disruptive levels. Shown in the image is latency.

into the scene. These are based on what the participants selected earlier in part two. Again, there is no interaction required from the participants.

The visualizations used in this user study were made straightforward to not complicate the study and not introduce more confounding variables. This includes showing the score and timer on the opposite stands in spatial relation, occasional graphics (target like the graphic on-field plane) appearing on the field, visualizing where and what event occurred, and a yellow downward arrow pointing at the referee of the game at all times. The arrow pointing at the referee simulates player tracking. However, we chose to use the referee as the referee was wearing a different color jersey, making the misalignment more distinct when there is an offset to the arrow so that participants will not get confused with the other surrounding players. There are also occasional statistical visualizations, such as the bar chart shown in Figure 5.5.

A total of three pilot tests were conducted before the user study, which allowed us to test and make a few changes to the experimental procedure and the study parameters.

Figure 5.5: Screenshot taken out of phase 1 of the user study (optimal condition user experience rating) showing a situated 3D bar chart of the line-out statistic of both teams.

Initial parameters were chosen based on group discussions with other experts in the team. The initial parameter for latency was set at 0-48 frames (2 seconds), and the registration accuracy was set to 0cm - 500cm (5m), which translates to a 5% offset when compared to the pitch size of 100m x 68m. The registration accuracy offset was done on the x-axis only to better standardize the conditions among the participants. The jitter frequency parameter was set from 5 seconds (less frequent) to 0.5 seconds (more frequent). Each time the jitter function is called, a random number of jitter happens from one to six occurrences on either the x or y-axis. We heuristically set the jitter magnitude at 1m and experimented with other values discussed with our expert panel. We found 0.5m slightly too subtle and decided to stick with 1m in which the jitter amplitude was similar to the jitter we observe on standard image target tracking AR.

During the first pilot test, we identified an issue with the term noticeable, as changes in factors are immediately noticeable when the slider is manipulated. Therefore, in our second pilot, we redefined the noticeable value to be a noticeable value in terms of user experience. In the second pilot study, we noticed that the participant did not understand what they were manipulating because we did not inform them of what condition they were manipulating. This was also present in the first pilot study, and we decided to let users know what they are manipulating to reduce the possibility of participants guessing and producing unreliable results. The disruptive value of most factors was lower than the noticeable value for the second pilot study. Therefore, we explained each condition to the participant for our last pilot study and fine-tuned our parameters. From the first two pilot studies, we found that no pilot study participant goes beyond 300cm for the registration accuracy. The participant in the second pilot study also complained that it was hard to fine-tune the registration accuracy as the parameter range was too extensive. Therefore, for the final study, we changed the parameter of registration accuracy from 0m-5m to 0m-3m. This means that the offset is reduced from 5% to 3% relative to the pitch size.

The final parameters of the screen manipulation slider for each condition are as follows. For latency, participants could slide from the optimal condition (minimum achievable latency) to a delay of 48 frames. Since the video playback is in 24 frames per second, this totals to a maximum delay of 2 seconds, and each frame would be 41.66 ms. For both conditions of registration accuracy, the parameters change from optimal to the maximum of 300 cm (3m), changed from 500cm as indicated in the pilot study. In the worst case, the visualizations are offset from the origin by 3 meters in global metric coordinates. Please note that we used a correct, surveyed 3D model of the stadium for this purpose. Finally, the slider for both jitter conditions affects the frequency of the jittering, which ranges from a delay of 5 seconds (less frequent) to 0.5 seconds (frequent) per sequence. The magnitude of the jitter is set to 1m away from its original position. In each sequence, the jitter occurs in either the x or y-axis in a positive or negative way randomly while also jittering from a random number of one to six. That is, if the participant positions the slider to the middle, a sequence will run every 2 seconds, with each sequence jittering one to six times. From looking at the resulting visualization, we believe that this mimicked the jitter that one would normally get from AR applications quite well.

5.4.2 Apparatus

For this study, we recorded 360° videos of an actual rugby game in a stadium using an Insta One X 360° camera. We used the video as a dynamic video background on a skybox in the indirect AR prototype. We use a 3D model of the stadium to align the content with the 360° video, which is done manually to ensure a good and consistent fit. We then superimpose visualizations onto the environment with spatial relation simulating an actual scenario one would encounter in the stadium during a live game. We conducted the entire study using an iPhone X with some part of the questionnaire (factor rating) done on the device. Other questionnaires (UEQ and other questions) are done on paper. We also observed the behaviors of participants and recorded interesting observations on paper. We implemented the prototype using Unity[1] and Vuforia[2].

5.4.3 Participants

We recruited 20 participants from the university for this study, mostly through word of mouth. There are no pre-requisites; no sports knowledge was needed to rate the user experience. Participants we recruited are between 20 and 30 years old ($\bar{x} = 24$, $\sigma = 2.66$), consisting of 11 women and 9 men. Among all participants, two participants have been to a game in the stadium at least ten times, while two participants have never been to a game. Participants were given a chocolate bar as a token of appreciation.

5.4.4 Procedures

We first gave participants time to read the information sheet and then invited them to complete the demographic questionnaire and consent form. After a brief introduction to the study, we presented the participants with the first part of the study. They were asked to rate their user experience after exploring an optimal indirect AR experience of an actual rugby game with overlaid visualizations. The goal of the first phase is for participants to evaluate the user experience via the UEQ, which would be the baseline for comparison with the other condition.

[1]https://unity.com

[2]https://developer.vuforia.com

Figure 5.6: Our user study prototype showing the latency factor delayed by 24 frames (1 second) controlled by the slider at the bottom. Optimally, the arrow should point at the referee (in white outfit).

Upon finishing the UEQ, participants are then briefed on what they would need to do in the second part, where they go through five conditions of technical factors. We walk them through the abovementioned factors to avoid getting lost in what is happening and provide feedback. Participants slide a slider at the bottom of the screen for each of the five factors to manipulate the technical factor (Figure 5.6). They are required to manipulate the factors until they find a noticeable impact on their user experience, then press a button to be brought back to the condition in which they now try to identify the disruptive value. The same steps are repeated for all conditions in a randomized order. Participants then move to part 3, which brings the noticeable values together to provide an experience with the compounded factors. Participants then filled in a different UEQ upon completion of the experience and then were asked if they would still use the prototype if the experience resembled the noticeable value condition. We also invite the participants to mention other factors that they think would impact the user experience.

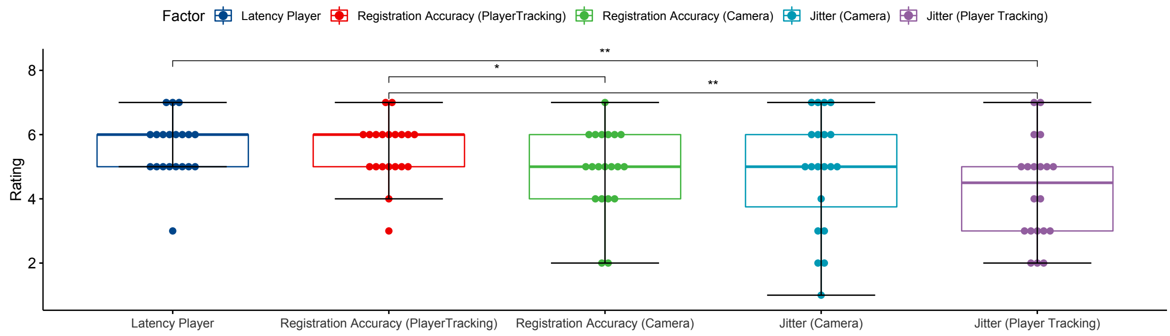


Figure 5.7: Impact Rating of Factors (on a 7 point Likert scale). Higher values indicate higher negative impacts on the user experience. Latency was ranked highest in impact on user experience, followed closely by registration accuracy of player tracking. Registration accuracy (Camera) and Jitter (Camera) seem to have a similar effect. Jitter (Player Tracking) received the lowest rating.

5.4.5 Results

We started by analyzing the results of the participant's ratings of the technical factors (7-point Likert scale). Our initial findings (Figure 5.7) show that latency ($\bar{x} = 5.6$, $\sigma = 0.94$) was rated highest having a negative impact on user experience, closely followed by registration accuracy (player tracking) ($\bar{x} = 5.5$, $\sigma = 0.95$), then Registration Accuracy (device camera) ($\bar{x} = 4.9$, $\sigma = 1.29$) and Jitter (device camera) ($\bar{x} = 4.85$, $\sigma = 1.81$). The Jitter factor (player tracking) was rated as having the smallest impact ($\bar{x} = 4.25$, $\sigma = 1.59$). We conducted a Friedman test to compare the ratings of the five technical factors. There, we found a significant effect of the factors on the rating (Friedman's chi-squared = 10.507, $df = 4$, $p = 0.033$). We then performed a Wilcoxon signed-rank test that indicated significant differences between latency (player tracking) and jitter (player tracking) ($p = 0.007$), between registration accuracy (player tracking) and jitter (player tracking) ($p = 0.007$), as well as between registration accuracy for device camera and player tracking ($p = 0.035$). We also analyzed Wilcoxon effect sizes r and found that there were large effect sizes between registration accuracy (player tracking) and jitter (player tracking) ($r = 0.621$) as well as between latency (player) and jitter (player tracking) ($r = 0.589$). We also found moderate effect sizes between latency (player) and registration accuracy (camera) (r=0.444, p=0.057), between latency player and jitter (camera) ($r = 0.327, p = 0.156$), between registration accuracy (player tracking) and registration accuracy (camera) ($r = 0.471, p = 0.035$) as well as between jitter (device camera) and jitter (player tracking) ($r = 0.382, p = 0.119$). The combination of small effect sizes of the remaining factors are latency (player tracking) and registration accuracy (player tracking) ($r = 0.023, p = 0.832$), registration accuracy (player tracking) and jitter (device camera) ($r = 0.243, p = 0.273$), registration accuracy (device camera) and jitter (player tracking) ($r = 0.238, p = 0.207$) and finally registration accuracy (device camera) and jitter (device camera) ($r = 0.038, p = 0.835$).

Interestingly, for registration accuracy, the offset of the dynamic content was rated to have a greater impact than the offset on the entire scene ($\bar{x} = 5.5$ compared to 4.9), despite the factor of the device camera that also affects the dynamic content. This is probably due to participants only focusing on the dynamic content during the player tracking registration accuracy factor; therefore, it makes a bigger impact than when everything is offset, and the participants focus on the overall picture. Participants found that the jittering of all visualizations was more disturbing ($\bar{x} = 4.85$) compared to when only the player tracking visualization was affected ($\bar{x} = 4.25$). In general, all of the tested technical factors do affect the user experience, as each one averaged above the midpoint of 3.5.

Analyzing and comparing the UEQ of part one and part three shows that the measurements are generally lower for the noticeable configuration compared to the optimal configuration, which is expected. The two-sample t-test (recommended in the UEQ handbook (Schrepp 2019)) also indicated that there are significant differences (p<0.05) between all components in the UEQ except for perspicuity (p=0.08). The other components are attractiveness ($p < 0.001$), efficiency (p<0.001), dependability ($p = 0.008$), stimulation ($p = 0.005$), and novelty (p=0.016). The scale means also shown in Figure 5.8. In terms of individual scores against the benchmark obtained from the UEQ Data Analytics Tool - version 4 (Figure 5.9), the optimal condition scored mostly above average and good except for dependability, which was on the border of below average. Meanwhile, the noticeable conditions scored mostly bad except for perspicuity and novelty, which still hovers in the below-average range.

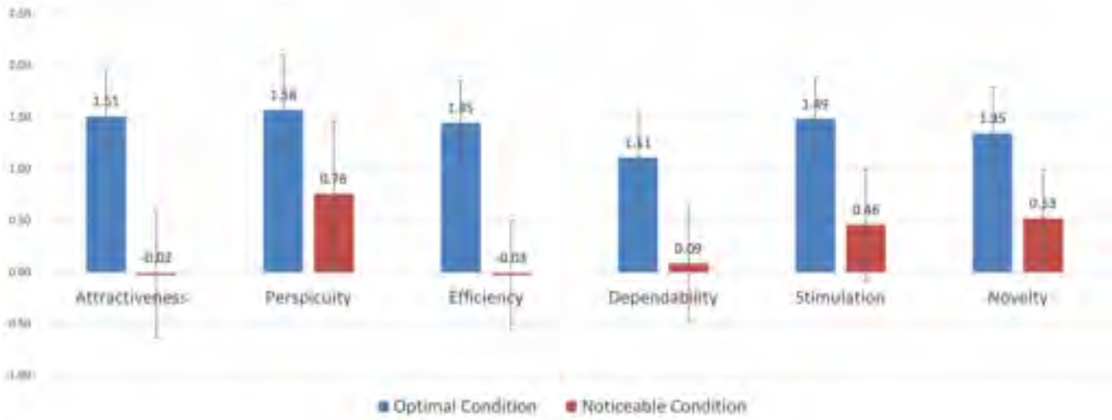


Figure 5.8: UEQ comparison of the optimal and the noticeable condition. The measurements for the noticeable configuration are overall lower compared to the optimal configuration. Our statistical analysis shows that there are significant differences between all components in the UEQ except for perspicuity.

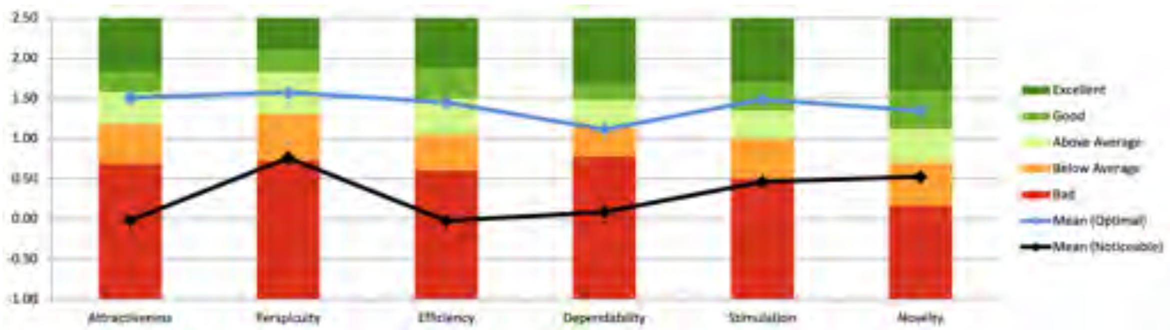


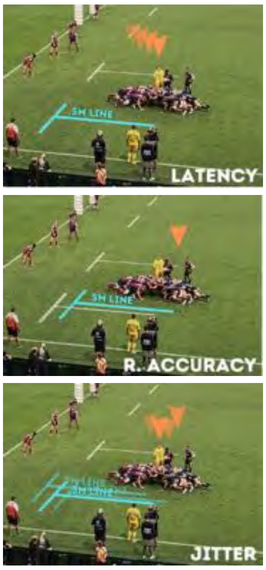
Figure 5.9: UEQ comparison between optimal configuration compared to noticeable configuration. Both curves are mapped to the original UEQ benchmark. As expected, the optimal configurations scored higher than the noticeable configurations. The scores for the optimal configurations are mostly at above average and good. The scores for the noticeable configuration were mostly in the flawed range.

In terms of consistency for both UEQ, both conditions have a Cronbachs Alpha of >0.7, with the noticeable condition having a higher average of the Cronbachs Alpha (0.87) compared to the optimal condition (0.81), showing high confidence in the scale results. In terms of inconsistencies, in both optimal and noticeable conditions, there is a participant (different participant in both conditions) with 3/6 categories where the answers are inconsistent with the scale (rating discrepancy for specific attributes in the same category). This is checked by seeing how much an item's best and worst evaluation on a scale differs. We decided that it was still acceptable and did not remove the data as outliers.

5.4.6 Discussion and Limitations

Answering **RQ4-1**, jitter was the lowest of the rated factors negatively impacting user experience, with a significant difference when compared to latency and registration accuracy for player tracking. Participants did not seem to care too much that the referee's arrow tracking was jittering, compared to when the whole visualization was jittering. P9 did not even notice the visualizations were jittering in the device camera jitter condition. Registration accuracy for dynamic content played a relatively large role in the user experience; presumably, pointing at the wrong player would greatly impact the game's understanding. Latency, as predicted, has quite an impact on the user experience, considering the disruptive value still falling below one second.

The average of the noticeable and disruptive values is tabulated in Figure 5.10, which answers **RQ4-2**. Note that the noticeable value refers to the value of a certain factor in which it starts to impact the user experience. We emphasize this because every slight manipulation in the factors would be quite noticeable but might not impact the user experience. The results show that for the latency of the dynamic content (player tracking), participants, on average, felt that the latency of 373.27ms started negatively impacting their user experience. In contrast, the latency of 654.90ms is borderline unusable. The registration accuracy of the camera and player tracking was similar, coming at 61.52cm and 62.35cm, respectively, for the noticeable value, with 105.78cm and 101.86 as the disruptive value. Jitter from the camera and dynamic content is also



Technical Factor	Noticeable Value	Disruptive Value
Latency (Player Tracking)	373.27ms	654.90ms
Registration Accuracy (Device Camera)	61.52cm	105.78cm
Registration Accuracy (Player Tracking)	62.35cm	101.86cm
Jitter - Interval (Device Camera)	4.62s	3.00s
Jitter - Interval (Player Tracking)	4.50s	2.81s

Figure 5.10: Average results for each factor comparing the noticeable value with the disruptive value. The factors from Figure 5.2 are placed on the left side for easier comprehension. Device camera factors affect all visualizations while player tracking only affects dynamic content such as the arrow as shown.

pretty similar at 4.62s and 4.50s between each sequence for the noticeable value with 3.00s and 2.81s for the disruptive value.

We identified that although we only used the noticeable values from the client device in the comparison condition, it had a strong negative impact on the user experience. The disruption to the user experience could be said to be non-linear as UEQ ratings fell somewhat steeply to values that are considered bad user experience, answering our **RQ4-3**. This implies that although we need to separate the factors while researching them, they alone are insufficient to prove a successful user experience. Each factor's compounding effect increases the frustration participants felt, prompting them to give a lower rating. One participant mentioned that they felt the factors did not affect them too much when considered individually but collectively impacted their user experience significantly (P19).

We have other insights regarding this observation in the form of a study limitation. Since the 360° video clip used for the manipulation of technical factors differs from the ones used for comparison, there are confounding variables present regarding how much horizontal movement (player running from left to right or vice versa) is present in that clip. There are more horizontal movements in the comparison clip that might amplify the effect of latency and registration accuracy.

Another observation is that the distance of action from the field might be a confounding variable. The action in the video clip starts relatively close to the camera and then moves towards the opposite side of the field. Therefore, participants who think they understand what they are looking for might immediately evaluate their experiences based on the action near them for the disruptive value, which differs if the action is happening further away. This was evident with one participant (P17) getting almost all the disruptive values smaller than the noticeable value, while three others had one set wrong (P5, P15, P16). We observed that the participant did it very quickly for the disruptive value, strengthening the earlier mentioned point.

Regarding additional information unrelated to technical factors, at least six participants complained or showed signs of arm fatigue from holding the mobile device. This prompted us to reconsider the ergonomics of using mobile AR during such a study, as a participant (P2) mentioned that any session longer than a minute started to introduce fatigue. In addition, three participants (P4, P5 and P9) had visible issues trying to manipulate the slider. This implies a usability issue with the user interface and could probably be solved by increasing the size of the slider handle.

5.5 Conclusion

This thesis chapter aims to serve as a starting point for research on improving user experiences influenced by technical factors and limitations. Our study focuses only on specific aspects of the three technical factors — mainly *tracking delay* and *networking delay* in latency, *static errors* in registration accuracy, and *spatial jitter* for jitter. Since we cannot avoid having technical factors that affect the system, the best we can do is reduce them to an acceptable level. Technical factors should be considered in the design phase of any AR application. Our study shows that technical factors, although at a noticeable value when compounded together, still significantly impact user experience.

Among the factors, both the content's latency and the registration's accuracy needed to be kept low to provide a coherent user experience. Although jitter is still disruptive to the user experience, it is not as bad as latency and registration accuracy if it happens only occasionally. However, the results and values presented in this study are for a large-scale AR environment. Different AR use cases with different distances to the object of interest in the environment would have variable results. Although we did not investigate this further, we foresee that future research could look into deriving a formula to calculate the acceptable values for these factors based on distance. We also think that this study performed on stereo 360° videos might be beneficial in better understanding the relationship between these factors and depth. We believe this part of the research will fill the gap identified by many other researchers mentioning the lack of user evaluation in AR publications (Olsson & Salo 2011, Han et al. 2018, Kim et al. 2018), particularly when considering technical factors for user experiences.

Chapter 6

Context-aware Interfaces for ARSpectator

Contents

Visualization methods and advances in hardware development alone would not be sufficient for enjoyable sports viewing experience on-site. In the past few chapters, we have discussed the *ARSpectator* overview, prototypes, visualizations, and technical factors that affect the user experience. The final aspect we investigated is how users interact with *ARSpectator*. For all the previous components to work together, a suitable interface is required to display all relevant information. However, developing a seamless and useful interface is not straightforward. The *ARSpectator* interface involves connections between multiple components of the *3'C's Situated Visualization for on-site Sports Spectating Framework*, such as *game context, user personalization* and *sports content*.

From Chapter 4, we found that half the participants wanted a personalized visualization. No one solution would fit all users and all game scenarios. In a sports spectating context, both interests in sports and the progress of game events are the main contextual factors that influence user experience (Sun & May 2009), illustrating the importance of the user context and the context of the game. Hence, for this chapter, we set out to develop and investigate AR interfaces for *ARSpectator* and the relevant components needed to make this happen with a focus on AR HMDs. We want to focus on bringing the best user experience for spectators rather than efficiency and performance matrix, similar to our perspective on the technical factors.

6.1 Design Space for an XR Interface

Just as we need an interface to interact with computers, such as using physical buttons on a keyboard or touching via a touchscreen, we need a suitable mode of interaction with XR applications. The Window, Icon, Menu, and Pointer (WIMP) interface paradigm is now considered a part of daily life. However, it is unsuitable for XR applications due to its design limitations for 2D spaces (Yang et al. 2019). Therefore, we would need to look into tangible user interfaces (Ishii 2008) or *Spatial Analytic Interfaces* (Ens & Irani 2016), where the use of 3D spaces introduces a closer relationship between virtual and physical objects.

6.1.1 XR Interaction Methods

Due to the casual nature of sports spectating, we focus primarily on interaction methods that would require the least amount of workload. Previous research has shown that interactions in XR can affect mental workload, especially for AR (Jost et al. 2020, Xi et al. 2022), which we want to avoid when a spectator goes to a game as they are there to enjoy themselves. We also would like to support interaction in large-environment AR scenarios, where the data referents (object semantically associated with the data) may be out of reach. This situation is contrary to the majority of AR applications involving a small AR workspace (Klein & Murray 2007) where the referents are close to the users.

Input Device

Most XR applications use controllers in combination with tailored interaction methods. A VR headset's tracked controllers provide button and trigger input and a mode of pointing. In a sports spectating scenario, it is not realistic for spectators to carry controllers such as those shipped with VR headsets, nor is it feasible for spectators to make large gestures with a controller during a game (Ahlström et al. 2014). A feasible solution is to use mini Bluetooth controllers with only one or two buttons for the essential interactions.

Voice Control

The use of voice to issue commands and control an interface is becoming quite common (Tulshan & Dhage 2018). However, it has limitations, such as the comprehensibility of the language itself (Dasgupta et al. 2018). Also, despite advances in noise cancellation (Lopez-Caudana 2011), voice control is not feasible in loud environments such as the stadium in our case, as the noise from other spectators is of similar frequency to the users' voice. Last but not least, voice input in any public setting could lead to awkward and embarrassing situations (Carter et al. 2015).

Gesture Control

There are two common forms of gesture control, the first involving gestures on the surface of a touch screen or trackpad and the other using the motion of hands in mid-air. We focus here on mid-air gestures since introducing touch-pad components may distract users' attention from the event they are spectating (Melax et al. 2013, Zhang, Bazarevsky, Vakunov, Tkachenka, Sung, Chang & Grundmann 2020). Some mobile phones and smart home devices use infrared sensors to recognize the waving of hands; however, we needed higher accuracy in XR applications. In XR applications, the whole palm, and not only fingertips, would need to be tracked, usually by the HMD's externalfacing cameras. There are also input devices such as data gloves that users could wear to track hands (Temoche et al. 2012), allowing for specific hand gestures to point and interact with virtual elements. Using such gestures in a stadium environment would be inappropriate as it quickly leads to arm fatigue for long-term use and might distract surrounding viewers, let alone potential social acceptance issues.

Head and Eye Tracking

Head and eye tracking both take into account where the user is looking. Head tracking (LaValle et al. 2014) is essential in all HMD, regardless of whether AR or VR. Both have been used for object selection; however, head tracking only approximates the field of view of the user based on their head movement, while eye-tracking (Clay et al. 2019) tracks the users' eyes and tries to pinpoint where the user is looking accurately. Eye-tracking is more easily implemented in VR systems as the camera is located close to the display lens, which is more difficult for an AR HMD. While eye tracking might be more useful in our AR sports spectating scenario, head tracking is still the more common form of tracking in many pre-existing devices. It would be the closest alternative to ascertaining a user's viewing direction.

Other Body Measurements

Physiological measurements such as electroencephalogram (EEG) (Mercier-Ganady et al. 2014), heart rate (Tumler et al. 2008), and galvanic skin response (Ventura & Porfiri 2020) could be used as feedback mechanisms toward what the user is feeling or wanting to do. An example of this would be the use of EEG to trigger input to the computer, known as a Brain-Computer Interface (BCI). BCI is meant to be used as a supplementary tool with other forms of interaction. With that said, there are still many limitations to such technology where users need to be still, and it would not make sense to rig spectators in the stadium with expensive and sensitive equipment.

6.1.2 XR Interfaces

After discussing the different input methods for XR, the next step is to determine how these input methods could be used or combined with other technologies related to XR into one interface. One can implement certain general design principles in AR, such as Shneiderman's design principles for direct manipulation (Shneiderman et al. 1993). However, with the added complexity of a 3D interface, we would need a better interaction interface. Here, we discuss some of the more common XR interfaces in other work.

Personal Interaction Panel

The Personal Interaction Panel (PIP) is a private panel used to interact with the contents of the user's surrounding environment. The original concept of PIP was presented by Szalavári & Gervautz (1997), where a panel held on the non-dominant hand is interacted with using a stylus on the dominant hand. This method of interaction allows users to manipulate virtual objects in the environment by performing various gestures. Using an AR or VR HMD, users could have visuals augmented on the PIP itself to provide context while providing privacy (Schmalstieg et al. 2002).

We can implement this interaction method in two different approaches. As mentioned, the PIP could be a physical panel the user holds, with the user interface overlaid virtually on the panel. This approach would give the user something to hold on to and is suitable for conditions where the user needs to perform specific movements such as minor adjustments with the stylus. The other approach would be to have a virtual panel. Simulating a physical PIP, a virtual panel could float in mid-air, freeing one of the users' hands while still allowing users to interact with the virtual objects in the environment through the virtual PIP.

Image Targets

It is common to see image targets in tangible AR applications, where a tangible user interface is bundled with an AR display to create a tangible physical interface (Billinghurst et al. 2005). As the name implies, an image target is an image that the camera of an AR system detects and identifies. The AR headset then tracks the movement of the image target and uses this as an anchor to place visualizations. We can use image target as an interaction method where images/codes printed on cards serve as user interface elements by moving the cards in a certain way or bringing image targets together. This is useful for applications that involve moving visualizations around, such as city planning, where the image targets could, e.g., represent buildings or chemical elements (e.g. Arloon [1]).

6.2 State Inference Model for a Context-Aware Interface

Context is the implicit situational information that occurs when humans communicate. However, this information does not transfer well to the interaction between computers and humans in traditional computing(Abowd et al. 1999). Therefore, to reduce the complexity of user input towards a system, the system needs to know the current information regarding the user environment, leading to increased research on context-aware computing (Liu et al. 2011). The increasing sophistication of the recently developed hardware allows better context-sensitive computing in sports (Baca et al. 2009). These include the advancement of sensors, cameras, computer vision, and deep learning technologies (Buric et al. 2019, Zhang, Wu, Yang, Wu, Chen & Xu 2020), which could contribute to detecting the context of the environment.

There are many parameters to consider when developing an interface with context awareness. Recently, a proposed model of enjoyment of sports spectatorship was introduced that emphasized the importance of context-awareness (Rogers et al. 2017). This model is based on self-determination theory (Deci & Ryan 2012), where competence, autonomy, and relatedness are considered intrinsic needs of people. This shows that we need to integrate the three sub-components of the *context* component in the *3'C's Situated Visualization for on-site Sports Spectating Framework* — object, user, and scene context (Chapter 4) for maximum effect. Objects refer to elements on the field, such

[1] http://www.arloon.com/

as players and balls, users are spectators themselves, and the scene is the collection of events happening in the game. This section will describe our state-inference model to enable a context-aware sports spectating interface. A context-aware interface would mean that spectators should get the right visualizations at the appropriate time and place without direct (manual) trigger events from users, which would distract from the enjoyment of the game to be followed at the same time.

Although we did conduct a user study with a specific interaction method which will be explained in the next section, our state-inference model is generic in terms of its interaction methods. Different instances of the state-inference model could use different interaction methods, from eye-tracking for gaze detection, head tracking, VR controllers, EEG, etc., as input methods. Hence, in this section we will not mention specific interaction methods and will use the terms *triggered* and *focused*.

6.2.1 State Inferences

We base the state inference model on the related work of Tsai & Huang (2018), proposing a user-behavior-driven augmented content display called iDisplay. The authors created a state inference mechanism that takes information from the users' movement and predicts user states based on historical user information. Therefore, based on different user actions such as walking, stationary or looking around, different visualization styles are presented based on what the system *thinks* the users' state is.

The core concept described echos closely what happens in the stadium, albeit our stadium use case is slightly different, where we rely more on the scene context to determine what is happening on the field. In the stadium, the sports data provider would provide the context of the object and scene, where player data (*object context*) and events that occur (*scene context*) would be streamed to the AR sports spectating system. The system will also take in *user context* in the form of user localization and personalization for determining the user's position in the stadium and what the user would be interested in seeing. An overview of the context-aware sports spectating state inference model is illustrated in Figure 6.1.

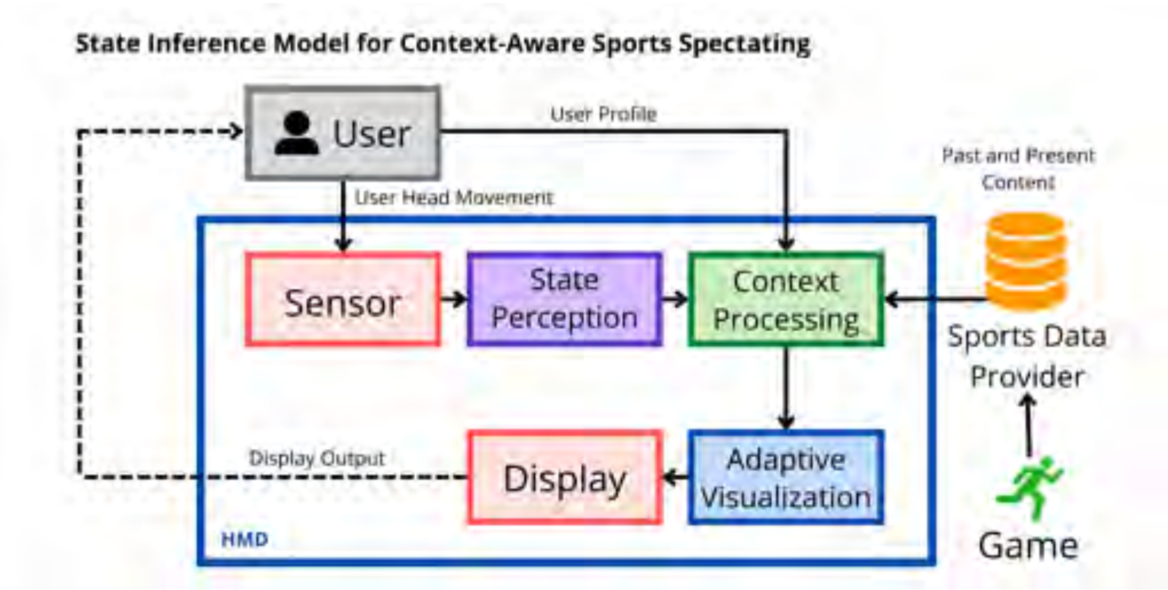


Figure 6.1: Overview of our state inference model based on Tsai & Huang (2018) which shows how the relevant visualization gets displayed to spectators based on the user, object, and scene context.

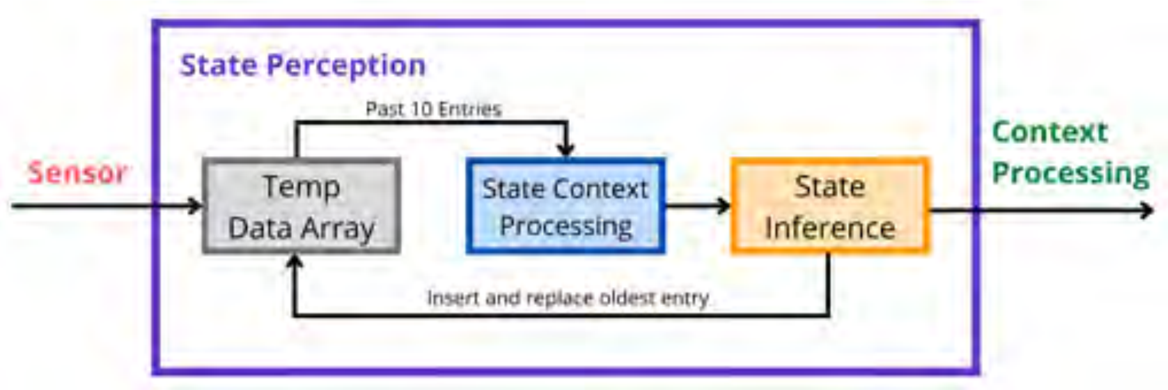


Figure 6.2: A closer look at the state perception part of the overall state inference model

Spectator State

The user provides two primary forms of context in the state inference model. First, closing in on the state perception part, the users' head movement in combination with the user's localization helps to determine the users' state (Figure 6.2) and with this to start identifying which canvas is in the field of view of the user. The second is the user profile, which could be ascertained with a pre-game questionnaire or social media profiles.

Using the onboard gyroscope sensors on the AR HMD and the assignment of "looking zones" in the stadium, we can identify where the users are looking in the stadium. It is also possible to calculate the head rotation speed to determine how fast a user looks around. These data then enter a temporary data array where, in our case, the past ten entries taken with an interval of 0.1s are used to determine what user state the spectator is in. We have defined four user states, although we acknowledge that there are probably more than four user states that a spectator will encounter in a real game. The four states are **Focused**, **Information Gathering**, **Crowd Exploration**, and **Agnostic**.

Therefore, if the system detects that the user is only staring at one area (low head movement and lingering on a specific zone), it can infer that the user is probably focused on something. Hence, the focused state will be triggered. If the spectator is looking around, there will be a change in head movement speed and zones the spectators are looking at, leading to an information-gathering state. If the spectator is focused or panning slowly at the crowd, it is possible that the user is interested in what others are thinking or might want more engagement with the crowd. This action will prompt the crowd-exploring state, which will display visualizations related to the crowd. While the state inference component then sends off the data to be processed with other contexts, a copy of it is inserted back into the temporary data array for the subsequent round of processing and replaces the oldest entry.

Game State

While a spectator state is individualized to each user, the game state is the same for everyone who attends that game. The game state's detection depends on an external sports data provider, which provides scene context to the system. For our prototype, we identified three general game states — **events**, **uneventful**, and **break time**. We could break down the events into major events, in which a scoring event occurs and minor events, such as fouls and substitutions. Characterizing all these game states allows us to display certain time-sensitive visualizations to the spectators and provide additional information to improve game understanding during uneventful times. During break time, if the spectator is still in their seat, it is possible to show a full on-field situated visualization, which would not be meaningful during a game due to the overlapping of visualizations and players on the field. The sports data provider also provides the object context when needed. This scenario is where certain player data could be displayed if, for example, the player scores or makes a foul play. If ball tracking is available, ball-path visualizations could be integrated into the visualizations as part of the object context.

Level of Detail

The Level of Detail (LOD) concept was often used in computer graphics methods to reduce polygons and rendering time (Heok & Daman 2004). However, this concept is also an integral part of sports spectating visualizations in general. A low LOD would show only some essential information when applied to situated visualizations, while more details are provided with a higher LOD. Since sports spectators are likely to focus on the game happening on the field, it is essential not to overload spectators with information when it is unnecessary. By introducing the LOD visualization method, spectators could get basic information about a specific game aspect and get more detail if they desire to do so. Otherwise, the LOD will not increase and hence prevent visual clutter. Visualizations then disappear after a while or when they are no longer needed.

For our sports spectating scenario, we developed four levels of detail for most of our visualizations. We decided to have four levels as we felt that it has the right balance between being showing too little at a time and overwhelming users with information.

- LOD-0: This is simply a small icon or text to notify spectators when visualizations are available. It is used when spectators are presented with the choice of what they want to see. (Top right and bottom right of Figure 6.3)
- LOD-1: The simplest form of visualization that involves data presentation. LOD-1 usually takes a graphical form and will not contain any form of text description. (Top left of Figure 6.3)
- LOD-2: This takes LOD-1 a step further and introduces some text elements to help describe some of the statistics shown. LOD-2 usually does not stray too far away from what LOD-1 looks like.
- LOD-3: Here is where the full details of the visualization are shown. It includes smaller text, such as numbers and percentages, where the information becomes specific. It could also contain additional data to supplement the LOD-2 visualization and, therefore, could be quite visually distinct from the previous LODs. (Bottom left of Figure 6.3)

Context Processing

Here, we describe the relationship between the spectator and the game states. The context processing step involves taking into account the spectator state, the game state, and the user profile to determine what visualization should be presented. Besides the visualization itself, it is also here where the system manages the LOD of each visualization, deciding if more detail is required for the spectator.

We illustrated how all the components work together with a state transition diagram (Figure 6.4). There are two methods by which visualizations could be triggered. The first method is when the system detects that the user is trying to gather information or explore the crowd, which is triggered by the user surveying the environment or looking at the crowd. At this stage, the LOD-0 visualizations would appear, and spectators

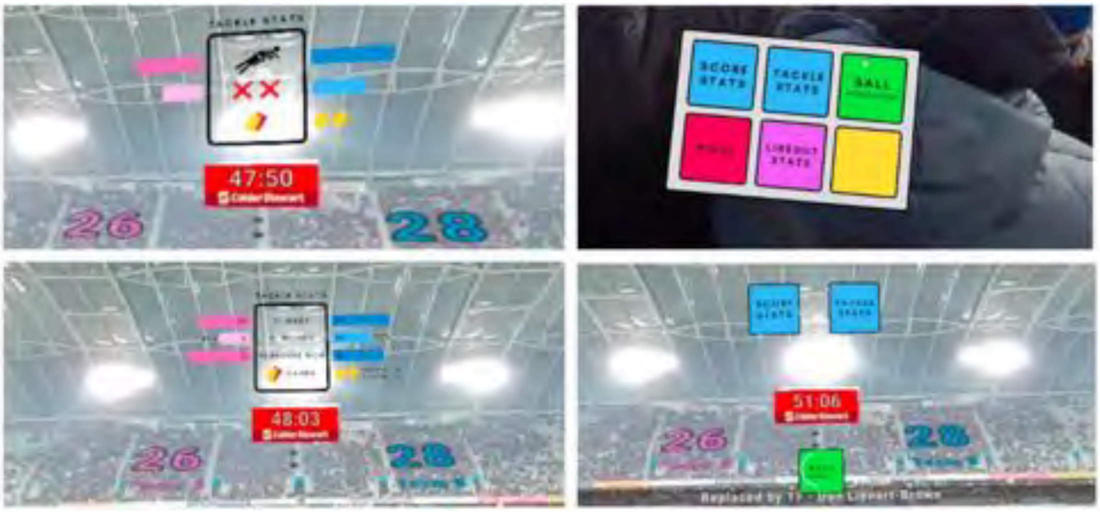


Figure 6.3: Our level of detail (LOD) implementation and two exemplary AR interfaces for sports spectating: Top Left) A LOD-1 rugby tackle stats visualization which displays minimal information. Bottom Left) A LOD-3 rugby tackle stats visualization with the maximum information shown if the user wants more details. Top Right) The manual trigger interface with LOD-0 visualization where users use a small controller to interact with visualizations. Bottom Right) Part of the *adaptive interface* where users rely fully on context and head movement to interact. Shown is the LOD-0 icons for user to select.

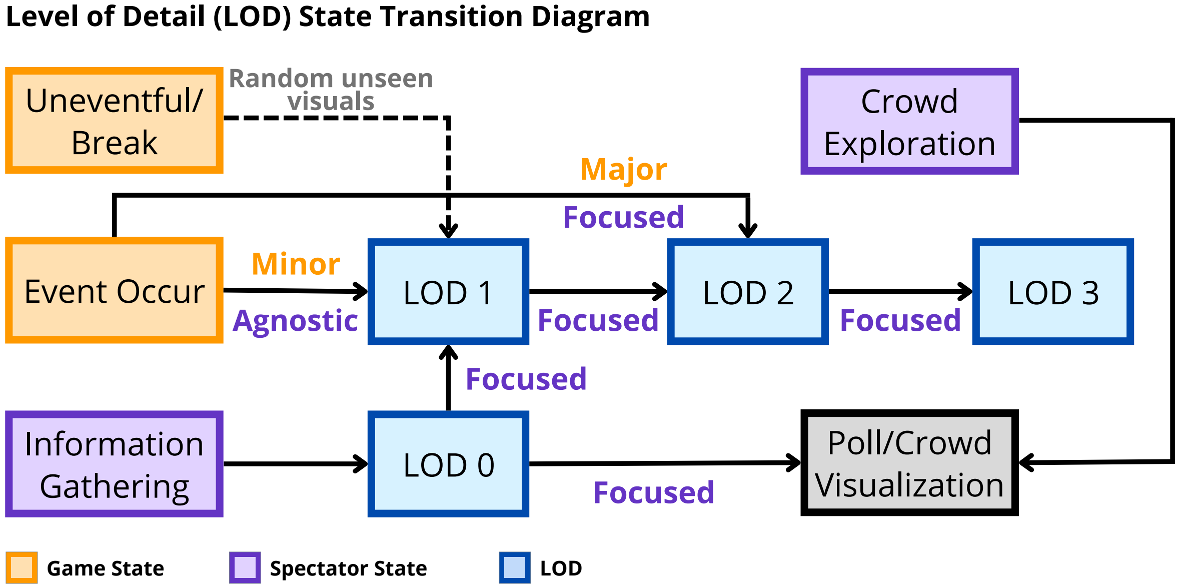


Figure 6.4: State transition diagram of the different LOD. Orange text are game-based and purple text are spectator-based states.

could trigger specific visualizations if they wish to do so. By activating a visualization, the LOD-1 visualization will be shown. If the user is exploring the crowd, a crowd-related visualization such as a poll will appear.

The second method for triggering visualizations is through the identification of the game state, where the context processing starts. While during major events, it is pretty clear what we should show to the spectators, during an uneventful time, the system should decide if a particular visualization that the spectator has not seen and might be interested in should be shown. This step is rather important as showing irrelevant content that spectators might not be interested in might distract them from the actual game rather than provide value. Usually, visualizations would take the form of LOD-1 in such cases; however, when coupled with other considerations such as event significance and user profile, it is possible to start directly with the LOD-2 visualization.

Regardless of which method triggered the LOD-1/LOD-2 visualizations (via information gathering spectator state or events in the game state), in order to increase the LOD, spectators would need to focus on the visualization. If the state perception determines that the spectator is focused, the visualization should evolve into the next

LOD after a couple of seconds. Hence, this gives spectators a choice; if they do not focus on the visualization, it will disappear after a while. After getting to LOD-3, the visualizations will stay there if the spectators are still focused on it (it will take more time to consume such detailed content) and disappear after a while if the spectators are no longer focusing on the visualization.

6.3 Implementing the XR Interface

Based on our experience and observations in prior games, there are three main points to take note of in a sports spectating experience, which form our requirements for an XR sports spectating interface:

1. Spectators are there to have a good time. Whether experienced sports spectators or casual spectators, everyone's end goal is to have an **enjoyable experience**, therefore, the interface should support this goal.
2. The stadium environment is **noisy and messy**. Unlike watching a movie in the cinema, the stadium environment is filled with life, from shouting sports fans to physical movements such as cheering and consuming food and beverages. The interface developed must adapt to such an environment.
3. **Interaction between fans** such as cheering, discussions, and celebrations are vital to a good stadium atmosphere. Therefore, the interface created should still allow fans to communicate with each other or even foster better interaction among the crowd.

Addressing these requirements, the XR interface for sports spectating should be unobtrusive, where spectators could still enjoy the stadium's atmosphere. Rather than being the most efficient method of information retrieval, the interface's focus would be to provide an enjoyable experience that complements the viewing experience. It would also need to withstand and adapt to the highly unpredictable nature of a sports game where for instance, fans could suddenly erupt in celebration.

To not disrupt the viewing experience, the interface would need to show information only when needed and at proper canvases in the stadium. This measure prevents sports spectators from getting distracted from the action on the field. The noisy and messy stadium environment also prevents us from implementing voice-based controls and other body measurements since spectators might be screaming and moving around. We also excluded gesture controls as spectators watching the sports most likely would not want to be distracted by the gestures of neighboring spectators. In addition, there is a potentially high cognitive load needed to perform hand gestures. Image targets are also unsuitable for our use case as most actions are happening far from the field, and it would not make sense for sports spectators to carry around image target cards.

More promising are considerations around the head and eye tracking, controllers, and the use of virtual PIP. We decided that since head tracking is more commonly available in existing devices, we would go with head tracking for now. We opted for a semi-transparent dot in the center of the screen, acting as a reticle (on-screen cursor). As for controllers, it would not be convenient for on-site spectators to carry around a bulky controller such as those used with VR headsets. If used at all, we aim to use a small controller with minimal control, such as the Hololens' clicker. Taking everything into consideration, as part of the ARSpectator experience, we implemented two interfaces to view visualizations in the stadium — a *manual trigger interface* and an *adaptive interface*.

6.3.1 *Manual Trigger Interface*

The *manual trigger interface* takes inspiration from the PIP, where the user would have their virtual interface panel that no one else could see. The panel is conveniently located to the bottom left of the spectator's view in the touch UI zone (Alger 2015), where spectators could easily reach without much eye strain. This placement is implemented as a world-referenced presentation canvas where the control panel is always anchored to the seat or a body-referenced canvas anchored to the users' body, acting as a belt. The latter would make more sense in a stadium, as users could constantly access their PIP by simply looking slightly down.

The content on the PIP should be the different visualizations that a user could trigger via a combination of head tracking and a one-button interaction device. The user needs to look down to the visualization they want to trigger and press the button on their device to display that respective visualization. When looking at the visualization shown in the stadium, users cycle through the different LOD with their button to a point where it turns off itself. The user then triggers the visualization they want as often as they like throughout the game.

6.3.2 *Adaptive Interface*

The *adaptive interface*, on the other hand, is an entirely hands-free interface. Using the same head-tracking interaction method as the *manual trigger interface*, users interact with the virtual elements by looking at a visualization for two seconds to interact. Visualizations, by default, appear depending on the game context, which factors in the ongoing event happening on the field, as well as the users' head movement.

For the game context, if there is, e.g. a penalty in the game, the penalty-related visualization might appear. At the same time, if there is not much action, some random visualizations that users have not explored from the visualization pool will appear such as tackle stats or voting polls.. This visualization pattern is similar to how sports broadcast feed visualizations to viewers. For the users' head movement, if the interface detects that users are looking for information, it will automatically spawn a menu that allows users to select what they want to see by simply looking at the various visualization options.

The visualization options of the *adaptive interface* are identical to the *manual trigger interface*, except when selecting a visualization, the option to select the visualization would appear on the canvas that the visualization is anchored. For example, suppose the score summary is to appear above the VIP seating box. In that case, there will be a co-located toggle, reducing the need for spectators to look down to select a visualization. A ring will circle the cursor dot to indicate that the user is selecting an option. The visualizations also allow for cycling through different LOD simply by looking at them. If the user is not interested in the visualization and does not look at it, it will disappear after five seconds without proceeding to the next LOD.

6.4 User Evaluation

Creating a fully functional and developed interface requires a complete cycle of iterative engineering. We developed an initial prototype to do a preliminary evaluation of the two interfaces described above to gain exploratory insight. Our goal is to evaluate the interfaces' user experience and get feedback on which interface users might prefer; getting feedback early in the development process is a crucial part of the agile methodology (Moniruzzaman & Hossain 2013). We decided to conduct a user study where participants could compare both methods and provide feedback while helping us understand what they want or seek in such an application. This study has received ethical approval (D22/099) from the University of Otago's ethics committee and followed enhanced pandemic hygiene procedures such as using hand sanitizers and disposable VR cover hygiene straps.

6.4.1 Dimensions of Interest

We are interested in the following dimensions of interest and assumptions:

1. Since both manual-trigger interface and adaptive interface are novel for the world of sports spectating; we anticipate that the participants will welcome them. Therefore, we expect the perceived user experience rating to be above average for both interfaces.

2. The LOD method for visualizing data is perceived as beneficial for the participants: depending on the information they need, they get the appropriate information.

3. The *adaptive interface* shows the correct visualizations with suitable perceived temporal and spatial accuracy.

6.4.2 Design and Apparatus

The design of this study is inspired by a study by Lindlbauer et al. (2019) in which the authors investigated an optimization-based approach in MR applications. They were interested in displaying not just the visualizations on the suitable canvases but also the correct LOD that users would need in that specific context. As the tasks provided to the participants change and the cognitive load of the user shifts, as measured by the Index of Pupillary Activity (IPA), the system will detect the context switch and proceed to manipulate the type of visualizations shown and the LOD for each of them. The authors used a VR environment to simulate an AR experience to eliminate the latency they experienced with a front (depth) camera.

We decided to reduce the complexity of the components of the study, namely the user modeling aspect and the canvas selection based on the FOV. During our informal exploratory interview with a sport expert, we found out that knowledgeable spectators would want to see more predictive analysis based on patterns from past matches, rather than simple box-score data or game statistics. With the limitations we have with data and resources, this is simply not feasible yet. Therefore, spectators are assumed to have a standard default spectator profile (casual viewer), and visualizations will appear in appropriate positions based on the visualization type rather than the participants' FOV. Instead of taking the task and cognitive load as input, we consider the state of the game and the state of the sports spectator. The system's output would be the visualizations with varying LOD, similar to the study referred to. We illustrate the comparison of both studies in Figure 6.5.

We designed a controlled laboratory study to explore the two interfaces we developed using our VR prototype, one of the indirect AR prototypes we created. Indirect AR prototypes are often a suitable method of simulating an AR experience in an off-site scenario (Wither et al. 2011), which was also done by the study we refer to (Lindlbauer et al. 2019). Our indirect AR prototype is created with the Unity game engine, using a pre-captured 360° video of an existing game to simulate the experience of a Rugby Union game. We then visualize the situated visualizations in the stadium environment as what an AR HMD would show, using the *3'C's Situated Visualization for on-site Sports Spectating Framework* from Chapter 4 as guidance. We informally interviewed an expert fan who explained the events in the 360° video snippet in Rugby Union for better accuracy and commentary of events. Both interfaces should be quite welcomed as there is no current way of visualizing information on-site. However, we are inter-

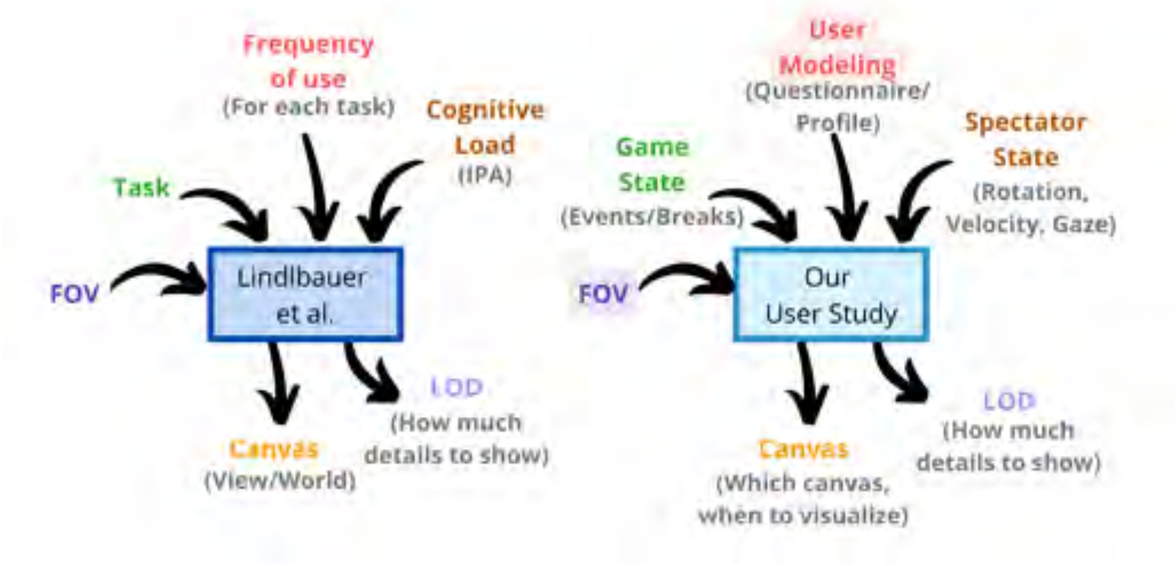


Figure 6.5: A comparison of the user study done by Lindlbauer et al. (2019) with our study. As a preliminary study to evaluate the interface, we simplified the study slightly by not including the user modeling and FOV component for this study (components highlighted).

ested to see which interface our participants prefer and its justification. Participants are only required to enjoy the game and summarise what they felt about the game dynamics. This provides something for the participants to focus on while evaluating the interfaces.

For both interfaces, the visualizations are the same. Some event-based visualizations appear to help describe what happened on the field. Those visualizations appear in-situ regardless of the interface. The interface only controls statistic-based visualizations. The statistics shown are made to be logical and coherent to all the other information provided and do not change throughout the experience. For example, the score shown on the stands would tally with the number of tries and penalties scored by each team. Visualizations are also anchored to canvases that spatially make sense to the data, such as putting the ball possession bar under the spectator stand for an easier understanding of which team has a higher possession rate. Most visualizations are placed in a way that does not interfere with the ongoing action on the field, except for the line-out statistics, which we place by the side of the playing field, as it is reasonable to place

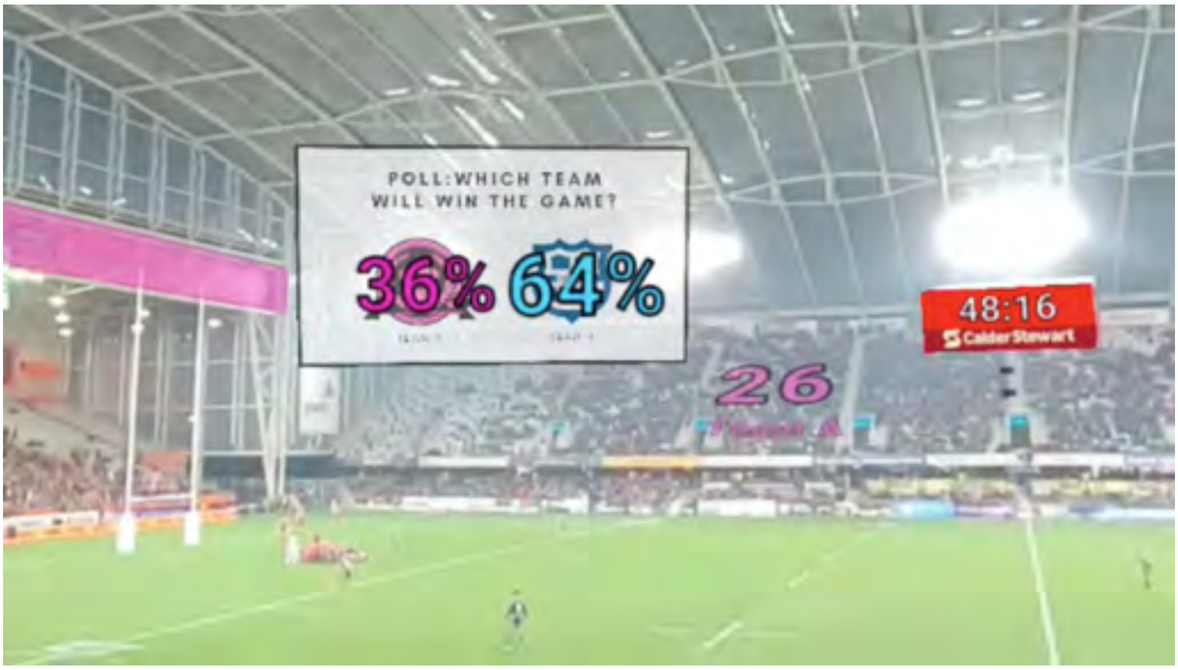


Figure 6.6: Screenshot of 360° video showing the poll visualization where spectators could vote for one of the two teams playing.

such visualizations by the line-out line.

In terms of hardware, participants experience the simulated AR experience in a Meta Quest 2 VR HMD, equipped with a VR cover and a disposable VR cover hygiene strap. This experience is a seated VR experience with no positional movement of the user or the 360° camera; hence, there would be a meager chance of any motion sickness occurring. The content of what the participants see is cast to a web browser on the researchers' computer for monitoring and troubleshooting purposes. The video used for the indirect AR video prototype is from 360° videos captured during a live rugby game from a static position in the stadium near the middle of the spectator standing on the long edge of the field (Figure 6.6).

6.4.3 Pilot Trial

We conducted a pilot trial with two participants to test the flow of the user study and to be alerted of potential challenges that might not have crossed our minds. Each of the pilot trial participants started with a different interface to obtain the perspective of the user study from both sides. Initially, apart from telling the participants the two different interfaces they would use, no additional instructions were provided. The first pilot participant started with the *adaptive interface* and was visibly confused with what was happening. Since the participant was looking around a lot, presumably just exploring the environment, the system determined that the participant was looking for visualizations and spawned the LOD-0 visualization too often. The participant was then confused about why the visualization kept appearing and did not seem to understand how the LOD system works. Upon finishing the *manual trigger interface*, the participant only got a grasp of the overall LOD system and gave feedback that it would be helpful if more information were given, as the mere two minutes of the video clip were insufficient for participants to understand what was happening.

For the second participant in the pilot trial, we added more detail to the method of interaction and how the *adaptive interface* works before the participant put on the headset. The second pilot participant started with the *manual trigger interface* first. In this case, the participant had a considerably better understanding of what was happening. However, in contrast to the first pilot participant, they did not move their head much during the experience. The experimenter gave a little prompt that if they required additional information — they could look around to get the LOD-0 visualization to spawn. Apart from that, the pilot study went well. We concluded that it would be better only to explain how the interface works when the participant has put on the headset so that they are less likely to forget. During the two pilot studies, we also tested some of the technical aspects of the study such as casting what the participant sees on the screen, optimizing the timing required for the study, as well as setting up a way to prevent the headset from going into standby mode, as that would stop the casting.

6.4.4 Participants

We recruited participants from the university through word of mouth and a mailing list of participants who participated in an unrelated AR study. In total, 20 participants (11 female and 9 male) aged between 18 and 37 ($\bar{x} = 25.25$, $\sigma = 6.26$) participated in our user study. Among the participants, only four participants have not been to a sports game in the stadium, while seven of them have only been once. Four participants have attended more than ten sports games in a stadium. All but two participants have experience with either AR or VR, and eight have experience with both.

6.4.5 Procedure

We first presented the participants with an information sheet. Upon reading it, we asked participants to complete the consent form and a demographic survey, and asked if they have symptoms of COVID-19. The participants are then briefed on the project's motivation and the user study's overall flow.

We told the participants that they will experience two different AR interfaces: the *manual trigger interface* and the other the *adaptive interface*. They are also briefed on the two interaction methods — (1) purely using head movement with a semi-transparent dot at the center of the screen as a cursor and (2) a combination of head movement and a press on the trigger of the controller.

We then give participants time to put on the headset and make sure they can see clearly and feel comfortable. The participants could then practice the interaction method by selecting their participant number (provided) from a menu and selecting the start button. This action will teleport the participants to the virtual stadium environment, where they will have time to look around and explore before the actual study begins. All participants were then briefed on the task they needed to do: watch the short two-minute clip and summarise what they felt regarding the game dynamics afterwards. The actual study conditions only start once they select the start button once more.

The conditions of the user study are randomized in a controlled manner. Depending on the allocated participant number, one of the two conditions will be presented to the participant when they press start again. For the *manual trigger interface*, the participant will see a menu panel on the bottom left of their view, simulating a panel attached to their waist (Figure 6.3, top right). This panel is the panel that controls the visualizations they see. For the *adaptive interface*, participants need to watch the

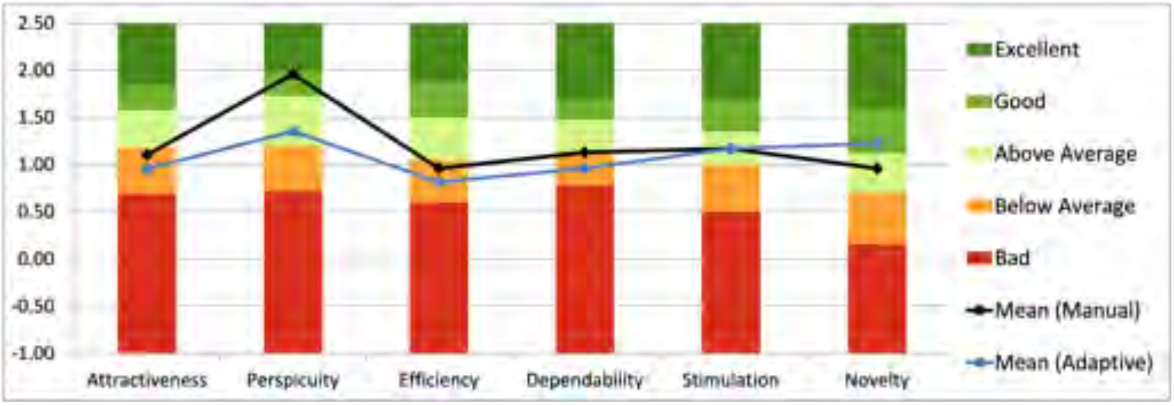


Figure 6.7: UEQ benchmark score (version 10) comparison of both AR interfaces, *manual trigger interface* (black graph) and *adaptive interface* (blue graph).

game, and visualizations will appear when necessary or if the system detects major head movement signaling that the participant is looking for additional information (Figure 6.3, bottom right).

The participants then experience the interface until the video snippet ends, which takes slightly more than two minutes for the clip. Upon finishing the first condition, we invite participants to complete a UEQ (Schrepp 2019) before proceeding to experience the second condition. The UEQ will allow us to know if the participants had a pleasant user experience using our prototype or not. The UEQ questionnaire consists of six scales (Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, and Novelty) with a range of -3 to +3, which is based on 26 responses the participants provide. After finishing the second condition, participants fill in another UEQ and follow up with a questionnaire that we created. The questionnaire consisted of questions regarding inconsistencies and distractions from Lindlbauer et al. (2019) along with some of our questions regarding appropriate visualization timing and placement.

After completing all questionnaires, participants are asked one last question about their take on the nature of sports spectating enjoyment in the stadium. They are then free to provide any additional feedback on the interface and study in general before a \$10 voucher is issued to them as a token of appreciation.

6.4.6 Results

In the UEQ data, we excluded a participant for being inconsistent in both UEQ measurements, with more than three of the six scales having inconsistent responses. The UEQ scored well with positive scales for both interfaces in all aspects. However, we note that the average scores of the *manual trigger interface* are higher in four aspects — attractiveness, perspicuity, efficiency, and dependability. This is compared to the *adaptive interface* except for stimulation and novelty. Despite this, no significant differences were found between the two conditions across all aspects when a two-sample t-test is conducted ($p = 0.71, p = 0.09, p = 0.67, p = 0.63, p = 1$, and $p = 0.29$). We also the UEQ of each interface against the latest benchmark provided by the UEQ Data Analytics Tool (version 10). Both interfaces scored one good score, two above averages, and three below averages, as shown in Figure 6.7. The results show that we can only partially confirm our first assumption about user experience ratings for both manual and *adaptive interface*.

For the additional questions, we found: For the question "I thought there was too much inconsistency in the system", participants rated $\bar{x} = 2.5$ and 3.05 for the manual and *adaptive interface*, respectively on a scale from 1 to 7 where 1 is "strongly disagree". The question "I find the interface and visualizations distracting" is slightly higher at $\bar{x} = 3.25$ and 3.65. These results show that participants are pretty satisfied with how the interface works, but they somewhat agree that it is slightly distracting. Most participants (13/20) preferred the *manual trigger interface* compared to the *adaptive interface*. Regarding the reason, P2, P9, and P16 mentioned the *adaptive interface* being distracting, while P5, P8, P9, P14, and P15 mentioned how the *manual trigger interface* has better control than the *adaptive interface*. P7, P11, P18, P19, and P20 all raised the issue of the wait time to select something for the *adaptive interface* and found this annoying. The feedback is supported by observations where for the *adaptive interface*, the information gathering context was triggered quite often, hence displaying the LOD-0 visualization for the participant regularly. However, we note that there might be a bias as most people will be used to triggers as they are widely used for computers.

We also asked participants to rate how "intelligent" they think the interface was in displaying the right visualizations at the right place and time for the *adaptive interface*. The ratings ranged once again from 1 to 7, with 1 meaning "very poorly". Participants rated $\bar{x} = 5.3$, 5.75, and 5.4, respectively, for the proper visualization, right place, and the right time. This result shows that the participants agreed that we had done the content, context, and canvas of the *adaptive interface* pretty well. P3 liked the poll option, deeming it nice to be involved in the game as a spectator. P4 recommended using sound to make the experience much more immersive, which we did not include to be able to communicate clearly with the participants. P4, P14, and P20 mentioned that head movement felt tiring or unnatural. P10 recommended making the *manual trigger interface* panel moveable so the participant could place the panel wherever desired. However, we believe that if we implemented eye-tracking, this would be less of a problem as participants do not need to look down so much with their heads.

Regarding observations, we found that almost all participants, especially P1, P3, P4, P13, P14, and P17, unintentionally triggered the adaptive mode's visualizations. This phenomenon is due to the amount of head movement done, probably because the participants want to explore more of the 360° video. However, this creates a distraction as the system keeps prompting the participants if they are interested in a specific visualization. P10 mentioned that it was more controllable in the *manual trigger interface*, and the participant did not know what was happening in the *adaptive interface*.

6.5 Discussion

Overall, from the results, we can summarize that the UEQs for both interfaces have no significant differences, with the *manual trigger interface* and *adaptive interface* performing well in different aspects. This section will discuss the implications of using AR in on-site sports spectating and what could be improved.

6.5.1 Active and Passive Roles in Sports Spectating

While sports spectators generally assume a passive role during a game, the use of an AR interface shifts the role of spectators to a more active role. On-site sports spectators could now consume content and participate in the scene. This participation does not just apply to the poll predictions we implemented but could also extend to crowd reactions (similar to social media emoji reactions). Spectators also now can view visualizations on-demand, creating different dynamics from what sports spectating traditionally is. However, as mentioned by P5 and P6, it could be potentially distracting. This feedback leads us to the discussion on the fine line between adaptive and distraction.

6.5.2 Fine Line Between Adaption and Distraction

We observed that many participants enjoyed the *adaptive interface*. Some participants (P1, P3, P7, P13, P17, P19) mentioned that they enjoyed the *adaptive interface* due to its hands-free nature. However, it could get obtrusive and distracting (P2 and P20) very quickly if the system is not well fine-tuned. In our user study, some participants were annoyed that they had to wait two seconds to move down the LOD in adaptive mode (P7 and P20). It was evident that some participants were puzzled when the system determined that the participant was looking for visualizations while they were looking around. However, this might be biased because the AR stadium environment is new to the participant, and they would not look around as much in reality.

6.5.3 Adaptive Manual Trigger Interface

From the previous discussion, it can be seen that the participants are keen to use an AR interface for on-site sports spectating. P5 recommended some form of context-awareness built into the *manual trigger interface*. We considered this and thought an adaptive *manual trigger interface* would be a good compromise between our two prototype interfaces. An adaptive *manual trigger interface* would provide some visual prompts at certain times, notifying spectators that there are updates to specific visu-

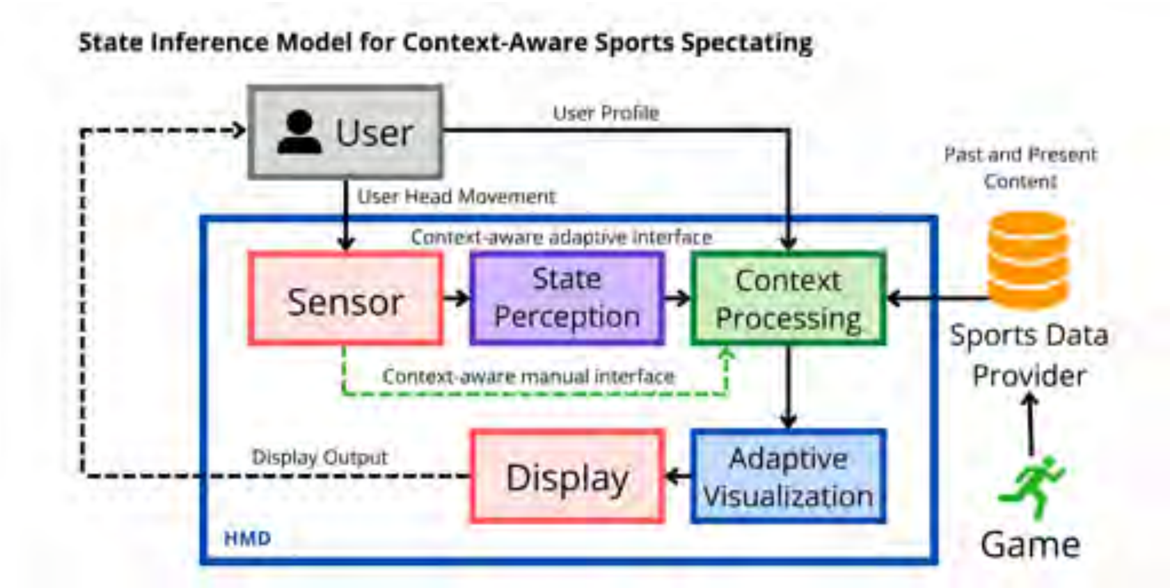


Figure 6.8: A modification to the state inference model to include a context-aware *manual trigger interface* (green dotted line).

alizations since they last visited, or some visualizations are recommended to be viewed now. However, whether to trigger the available visualizations is ultimately the user's choice. If the participant is very focused on the game and does not want any distractions, they could choose to see the visualizations early. We took the state inference model and made a slight modification to accommodate a context-aware *manual trigger interface* (Figure 6.8).

With a context-aware *manual trigger interface*, we could omit the state perception component from the process. This would mean that the system still considers the scene context (game state) and user context (spectator profile) except for the spectator state. Therefore, spectators have the freedom to see visualizations when they want to and will still be prompted about new or recommended visualizations without the distraction of unprompted visualizations. However, this approach has not yet been implemented and validated. Therefore, given the opportunity, we would like to implement and evaluate such an interface in the future.

6.6 Conclusion

In this chapter, we investigated how users could interact with *ARSpectator* as it plays a vital role in providing a good user experience. Based on the user study conducted in Chapter 4, we know that a personalized viewing experience is necessary. The nature of sports spectating also indicates that a simple hands-free interface is preferable for our use case. Therefore, considering the points mentioned, we proposed a state inference model for context-aware sports spectating. This model utilizes the LOD concept to present information in varying complexity, depending on the user context, as part of the discussion in Chapter 3. Using our state inference model, we developed two interfaces — the *manual trigger interface* and *adaptive interface* for comparison and evaluation.

This initial study provided a starting point to create context-aware interfaces in AR sports spectating. In summary, for our *adaptive interface* to be successful, it must be reasonably accurate in predicting what the user wants to see. This result mirrors the finding of Rogers et al. (2017) stating that fans will enjoy a sporting event more when they feel competent and autonomous. If the *adaptive interface* has not reached that level of competency, a manual approach is preferable to give the user more control. More work is needed to examine an adaptive *manual trigger interface* where the best of both worlds could be achieved before a fully fine-tuned *adaptive interface* is implemented.

Chapter 7

Conclusion

Contents

This thesis was written with the objective of working towards the use of AR in on-site sports spectating during live games. Sports broadcast viewers often have access to more information regarding the game due to the presence of a vision mixer[1] who annotates visualizations on the broadcast footage. This research aims to bridge the game understanding gap between sports broadcast viewers and on-site spectators using situated visualization through AR. This mimics sports broadcast viewers' visualizations while providing personalization and interaction.

Our prototype solution — *ARSpectator*, is an example of how we could achieve this by providing an overview of the system in Chapter 3. It consists of four main components: the *content source, user tracking, content aggregator*, and *mobile client*. In the same chapter, we proposed a *Flexible XR Prototyping Framework* to solve the challenge of developing an on-site AR application remotely and introduced four dif-

[1]A person who chooses what to show the broadcast audience, including visualizations made by the graphics operator.

ferent prototypes in our use case — a miniature lab AR prototype, a mobile indirect AR prototype, a VR prototype, and a stadium AR prototype. We found that the indirect AR and VR prototypes are helpful to developers in our off-site development and evaluation process, serving as a simulation of what users will see in the actual AR scenario.

We explore visualization methods in Chapter 4, where two novel situated visualization methods — *Situated Broadcast-styled Visualization (SBV)* and *Situated Infographics (SI)*, were developed to benefit viewers and provide information during the game through our proposed *3'C's Situated Visualization for on-site Sports Spectating Framework*. We found that the participants liked the visualization method due to the similarity in layout with sports broadcasting for *SBV* and the better spatial understanding in *SI*. Ultimately, we proposed a hybrid visualization method that combines the advantages of both methods while not cluttering the screen.

Since our research also focuses on the user experience, we investigated and found that latency and registration accuracy are the two technical factors that participants find affect their user experience more than jitter (Chapter 5). The other element that could disrupt the user experience is the ability of the context-aware interface, which is an interface that provides information to the user based on what is happening in the surrounding, including the state of the user. The state of the user helps identify what spectators want to see and provides a seamless experience if done well. However, failure to do so leads to distraction, and spectators would want manual control over what they see (Chapter 6).

The research then leads to this chapter, where we compile our summary of the findings and discuss what we learned during the research process. In the following sections, we explore our *Stadium of the Future* vision and discuss what future work could be done to achieve this goal.

7.1 Summary of Contributions

Here, we revisit the research questions and discuss our contributions to them. We also discuss the lessons we learned throughout the research process that we did not consider prior to the start of the research.

In **RQ1**, we were interested in knowing what components are needed for an on-site AR sports spectating use case and the interactions between the components. We found that four main components are necessary to provide spectators situated visualizations. The first component is **content sources**, which includes user-profiles and provides server-side game information. It then combines the users' **tracking data** to be compiled in the **content aggregator**. Finally, both the tracking data and content aggregator will end up in the **mobile client** where the visualizations are shown to the user.

RQ2 seeks to find out how to develop different XR prototypes flexibly to use in sports spectating. Our solution to this research question is the *Flexible XR Prototyping Framework*. Developers of AR applications should incorporate these three aspects into their planning — the **locality**, **scale**, and **evaluations** to justify why different XR prototypes are required. A modular design is then needed to standardize most scripts and prefabs across the various prototypes. This includes dynamically instantiating objects and using global coordinates to prevent complications in specific prototypes. Lastly, as part of the modularization, we have three components — the **tracking manager**, **visualization manager**, and the **data manager**. This framework is a good guideline for standardization throughout the different prototype modes.

There is no clear answer regarding what AR visualization approaches would allow for a more enjoyable on-site sports spectating experience, as mentioned in **RQ3**. The participants welcome both *SBV* and *SI* created through the *3'C's Situated Visualization for on-site Sports Spectating Framework*. Although it has not yet been evaluated, the answer might be a *hybrid* approach, where the advantages of both approaches are combined, making it the ideal visualization method. However, what is certain is that both *SBV* and *SI* are providing a positive impact on the user experience; hence, either method would be a good starting point.

We then examine **RQ4** on how technical factors in AR affect our large-scale environment on-site AR user experience. In order to answer this, we conducted a user study using mobile indirect AR where participants could manipulate three technical factors, which are latency, registration accuracy, and jitter. Participants get to choose at what level the factors noticeably affect the user experience and at what level the factors are disruptive. Our results indicate that latency is the number one factor affecting user experience in a sports spectating scenario, followed closely by registration accuracy. We also found that combining multiple noticeable values of factors greatly increases the disruption of the user experience.

We developed a manual trigger interface and an adaptive interface to provide game-related information to spectators without disrupting their sports spectating experience, as mentioned in **RQ5**. We had similar incoherent results to **RQ3** where the participants were divided on the answer and suggested a hybrid approach of both. We concluded that for a context-aware interface to work, it would need to be reasonably accurate in predicting what is of interest to the spectator. The line between being an intelligent adaptive system and a distraction to the game is relatively thin but has yet to be discovered. Until the adaptive interface is fine-tuned, despite many participants liking it, many would still prefer to have control over what they see.

During the implementation of the four user studies, we realized that sports spectators often want to have a good time and enjoy the game's overall experience. Therefore, it does not make much sense to have specific tasks for the user study participants to complete. Although we did have tasks for the first two user studies, we found that what we wanted to measure was the user experience, and the inclusion of the task might cause stress to the participants and does not reflect the reality of the use case. However, when we removed the tasks in user study 3, we were also concerned that participants might not have anything to focus on and wander off in their minds. We ended up with a compromise where the task for the user is to observe the game as if they were on-site in the stadium and give a summary of what they thought about the game. This method allows the participant to focus on something during the user study while not overloading them with tasks that require a lot of cognitive and mental workload, such as user studies 1 and 2.

7.2 The Stadium of the Future

Lately, the *Stadium of the Future* has received much attention, as shown in the number of online articles cited later in this paragraph. What could be agreed on is that *Stadium of the Future* would definitely be a technological fortress, packed with advanced hardware to provide a better experience to on-site spectators (Mons 2020, *Stadium of the Future* 2019). This is enabled by the willingness of sports teams to potentially invest USD 10 billion by 2030 in stadium infrastructure (Young 2022). Our vision on the *Stadium of the Future* mirrors the current trend, but we believe that XR technologies will play a significant role in it. In this section, we would like to go through a few points of how XR technologies could enhance the *Stadium of the Future*.

7.2.1 Interactive Viewing Experience

Our vision of the *Stadium of the Future* is in line with the primary motivation of this thesis — to bring AR experience on-site. We imagine that in the future, spectators will be able to get additional information on the field through situated visualization and AR HMDs. As of now, such AR HMDs do not exist yet, current HMDs such as the Microsoft Hololens 2 are still bulky and are not suitable for long term usage and large environment AR. As demonstrated in this thesis, we have already explored aspects of this vision. We predict that the XR technology will be widely available in the future, similar to how social media gained popularity (Greenwood et al. 2016).

In addition to that, it would be possible to have replays on-demand, similar to what Narasimhan et al. (2015) proposed but done in a situated approach on an AR HMD. To illustrate the concept, we did a mock-up placing players of a soccer video clip from Rematas et al. (2018) into our rugby stadium (Figure 7.1). The replay would work great during breaks using image inpainting to place virtual players on the empty field (Kari et al. 2021). However, it might also be possible to have replays on our PIP (personal interaction panel) as an interface, which we discussed in Chapter 6. Imagine if spectators could bring out a tracked board (similar to the size of a mobile tablet) which acts as their PIP with the field augmented on it. Spectators then could see players on their mini-field similarly to Figure 7.1 anytime during the game. They

Figure 7.1: We combined soccer players from Rematas et al. (2018) into our Rugby indirect AR prototype to provide our impression on how AR replays would look like.

could also manipulate the board to see the field from various perspectives.

From the same chapter, we also learned that most people go to the stadium to enjoy the stadium atmosphere and the community feel with their friends. We predict that in the *Stadium of the Future*, on-site spectators would be able to interact with the visualizations and with the other spectators around them. An idea that we had and illustrated in one of our AR prototypes is to have crowd reactions (Figure 7.2). Similar to social media, where people could react to posts, spectators could now react to events in a game, and it will be visualized with spatial relation to approximately where the spectator is. There are even eyeglasses created such as AffectiveWear (Masai et al. 2015) that recognizes facial expressions that allows aggregation of reactions to be done automatically (Kunze et al. 2017). This additional engagement would be a factor in attracting more people to the stadium.

7.2.2 Inclusive Viewing Experience

The usage of XR technology opens up opportunities to design a more inclusive *Stadium of the Future* environment for spectators with disabilities (Ryskeldiev et al. 2021). This

Figure 7.2: In one of our AR prototype, we simulated crowd reaction on the stands — a feature where spectators could react on how they are feeling at it would appear on the stands approximately where the spectator is seated.

is in line with the concept of augmented human (Papagiannis 2017) where technology is used to extend the limits of human capability. XR technology could enhance the experience of spectators with vision problems by providing visual and haptic assistance when needed, allowing them to feel the stadium's atmosphere while understanding what is happening. Spectators with hearing or speaking disabilities could also benefit from XR technologies in interpreting crowd atmosphere and expressing themselves, as mentioned in the previous section on virtual cheers and spectator augmentation (Kunze et al. 2017).

7.2.3 Hybrid Stadium Experience

The COVID-19 pandemic has changed the sports spectating scene, affecting both players and spectators (Wong et al. 2020). Some spectators are less inclined to return to the stadium and prefer to watch live broadcasts of such events (Gorman 2021). The pandemic also highlighted how we could use XR technologies to enhance remote experiences in the tourism and education sector (Shen et al. 2022).

Therefore, it is logical that companies will try to work towards bringing the on-site experience to remote viewers via immersive XR technologies. For example, our *VR indirect AR prototype* combined with a live broadcast would allow remote viewers to experience the on-site experience. With multiple 360 cameras placed at different locations in the stadium, we could even provide a perspective change for viewers, depending on what they want to see (Shishido et al. 2020). By reconstructing the whole game, we could also provide spectators replays of the game from a first-person perspective, similar to the work done by Ferrer et al. (2016) for enhanced training. Besides that, stereo 3D 360 cameras could potentially provide better immersion, something we captured with an Insta Pro 2 multi-lens 360 camera in an empty stadium but have not investigated further. Further integration with the on-site crowd participation component mentioned in the previous paragraph would also enhance both the on-site and remote viewing experience.

Besides the personal viewing experience, the other form of hybrid stadium experience would be similar to how some conferences hold satellite events. Satellite events are events hosted in multiple locations or venues, giving participants the flexibility to participate no matter where they are (Maybee 2021). Even without the pandemic, travel to the game venue is not always feasible or financially possible. Continuing from our replay idea mentioned in the previous section, we could hold mini-events where sports fans could gather in smaller groups. They could watch a game on broadcast with assistance from the *lab AR prototype* developed in Chapter 3 as a supplementary tool. This approach includes elements of spectator atmosphere while providing access to an exciting method of spectating a game — using a mobile device while watching the action on a tabletop.

7.2.4 On-site Community Events

In order to justify the extra cost involved in participating in an on-site sports game, we believe that future stadium events would incorporate local exhibitions and community events that are exclusive to on-site attendees. It could be a fans' meet-up or exhibition regarding the history of the sports team, similar to how academic conferences often have an accompanying industry exhibition. XR technologies could also be applied here, from large-scale XR such as projection-based AR implemented at Disneyland (Mine et al. 2012) to personalized XR experiences such as a mobile AR treasure hunt. Interactive installations such as those in museums (Huang et al. 2018) could also be done in the stadium exhibitions, providing a unique experience for attendees.

Another community activity is to take advantage of the physical space and implement location-based games (Fonseca et al. 2022) using XR technology. The games would promote focused social interaction (Bardis 1979), where everyone physically present has the common goal of enjoying the sport. It could be in the form of an on-site AR treasure hunt or games where supporters of the opposing team could play collaboratively to compete against each other. These provide additional value to attendees to justify the extra cost and incentive to attend a game physically.

7.3 Outlook

Our work sets the direction for future research in this field to advance toward the *Stadium of the Future* concept. There is still room for research regarding developing a polished application for on-site AR sports spectating. Even for the prototypes we created, we could make improvements, such as including a smaller scale stadium model for users to get a better idea of where they are. Danyluk et al. (2021) name this method having multiples for a Worlds-in-Miniature Interface, where a smaller scale of the actual visualization is shown for better context. Spectators could identify positioning based on a small-scale map of where they are in relation to the stadium.

During our brainstorming sessions, we were interested in integrating actual broadcast footage to view the game or replays from a different angle. We would also like to incorporate more crowd participation components into our prototype as spectators' participation could help improve the atmosphere of the event (Ludvigsen & Veerasawmy 2010, Centieiro et al. 2014). Similar research has been done for live sports broadcast (Centieiro et al. 2015), but we want to integrate this into the stadium by sharing spectator reactions. Extending our current scene context mentioned in Chapter 4, future work in this could include automated crowd "atmosphere detection", possibly by environmental context-aware sound detection. Future work could include exploring the potential effect of design factors on user experience, using similar techniques done in Chapter 5 on the contents of Chapter 4. Near-perfect system implementation will still lead to a sub-par user experience if the actual, overlaid visualizations are poorly designed or placed.

It is also necessary to simultaneously test the interfaces we developed in Chapter 6 for an extended period of game time with other participants. This is to test the social acceptance of our interfaces and the general use of such technology in social events. Schwind et al. (2018) observed that in situations where people are supposed to interact with each other, the social acceptability of using VR HMDs is reduced, which might be the same for AR HMDs. Due to the limitations of the COVID-19 pandemic, we struggled even to get enough live games where we could do longer high-quality recordings, not to mention conducting on-site studies. The duration of our video clips might give participants time to evaluate the user experience of our interface but is insufficient to see if the interface aided in game understanding or led to realizations of any particular insight.

Lastly, since we focus only on visualization and user experience in this thesis, many technical aspects of *ARSpectator* are still not discussed. Our evaluation of the technical factors affecting user experience in Chapter 5 covers only some of the subsets of the three technical factors. Factors such as *application delay*, *rendering delay*, *display delay*, *temporal jitter* could be studied in the future. We hope that our initial findings will help advance future research to investigate this in different use cases. This would build a better understanding of how optimized the technical system should be and if it is worth the cost for that extra millimeter of accuracy or precision. Apart from that, we would also need to consider the stadium network infrastructure, processing power of most participants' mobile devices, localization efficiency, and many more factors.

Acronyms

AR Augmented Reality

ARL Animation Research Limited

CAD Computer-aided Design

CAVE CAVE Automatic Virtual Environment

FOV Field-of-view

HMD Head-mounted display

IPQ Igroup Presence Questionnaire

LOD Level of detail

MR Mixed Reality

OST Optical see-through

PIP Personal Interaction Panel

POV Point-of-view

SBV Situated Broadcast-styled Visualization

SI Situated Infographics

SLAM Simultaneous Localization and Mapping

TI Traditional infographics

TLX Task Load Index

UEQ User Experience Questionnaire

VR Virtual Reality

VST Video see-through

XR Extended Reality

References

Abowd, G. D., Dey, A. K., Brown, P. J., Davies, N., Smith, M. & Steggles, P. (1999), Towards a better understanding of context and context-awareness, *in* ‘International symposium on handheld and ubiquitous computing’, Springer, pp. 304–307.

Adesida, Y., Papi, E. & McGregor, A. H. (2019), ‘Exploring the role of wearable technology in sport kinematics and kinetics: A systematic review’, *Sensors* **19**(7), 1597.

Afify, M. K. (2018), ‘The effect of the difference between infographic designing types (static vs animated) on developing visual learning designing skills and recognition of its elements and principles’, *International Journal of Emerging Technologies in Learning (iJET)* **13**(09), 204–223.

Agusanto, K., Li, L., Chuangui, Z. & Sing, N. W. (2003), Photorealistic rendering for augmented reality using environment illumination, *in* ‘The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.’, IEEE, pp. 208–216.

Ahlström, D., Hasan, K. & Irani, P. (2014), Are you comfortable doing that? Acceptance studies of around-device gestures in and for public settings, *in* ‘Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services’, pp. 193–202.

Akçayır, M., Akçayır, G., Pektaş, H. M. & Ocak, M. A. (2016), ‘Augmented reality in science laboratories: The effects of augmented reality on university students’ laboratory skills and attitudes toward science laboratories’, *Computers in Human*

Behavior **57**, 334–342.
URL: *https://www.sciencedirect.com/science/article/pii/S0747563215303253*

Alger, M. (2015), Visual design methods for virtual reality, Master's thesis.

Amin, D. & Govilkar, S. (2015), 'Comparative study of augmented reality SDKs', *International Journal on Computational Science & Applications* **5**(1), 11–26.

Apostolou, M. & Lambrianou, R. (2017), 'What motivates people to do and watch sports? Exploring the effect of sex, age, partner status, and parenthood', *Evolutionary Psychological Science* **3**(1), 20–33.

Azuma, R. (1993), 'Tracking requirements for augmented reality', *Communications of the ACM* **36**(7), 50–51.

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S. & MacIntyre, B. (2001), 'Recent advances in augmented reality', *IEEE computer graphics and applications* **21**(6), 34–47.

Azuma, R. & Bishop, G. (1994), Improving static and dynamic registration in an optical see-through HMD, *in* 'Proceedings of the 21st annual conference on Computer graphics and interactive techniques', pp. 197–204.

Azuma, R. T. (1997), 'A survey of augmented reality', *Presence: Teleoperators & Virtual Environments* **6**(4), 355–385.

Baca, A., Dabnichki, P., Heller, M. & Kornfeind, P. (2009), 'Ubiquitous computing in sports: A review and analysis', *Journal of Sports Sciences* **27**(12), 1335–1346.

Bach, B., Sicat, R., Beyer, J., Cordeil, M. & Pfister, H. (2017), 'The hologram in my hand: How effective is interactive exploration of 3D visualizations in immersive tangible augmented reality?', *IEEE transactions on visualization and computer graphics* **24**(1), 457–467.

Bach, B., Sicat, R., Pfister, H. & Quigley, A. (2017), Drawing into the AR-CANVAS: Designing embedded visualizations for augmented reality, *in* 'Workshop on Immersive Analytics, IEEE Vis'.

Baker, L., Ventura, J., Zollmann, S., Mills, S. & Langlotz, T. (2020), Splat: Spherical localization and tracking in large spaces, *in* '2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)', pp. 809–817.

Baker, L., Zollmann, S., Mills, S. & Langlotz, T. (2018), SoftPOSIT for Augmented Reality in complex environments: Limitations and challenges, *in* '2018 International Conference on Image and Vision Computing New Zealand (IVCNZ)', IEEE, pp. 1–6.

Barba, E., MacIntyre, B., Rouse, R. & Bolter, J. (2010), Thinking inside the box: making meaning in a handheld AR experience, *in* '2010 IEEE International Symposium on Mixed and Augmented Reality-Arts, Media, and Humanities', IEEE, pp. 19–26.

Bardis, P. D. (1979), 'Social interaction and social processes', *Social Science* **54**(3), 147–167.

Barsom, E. Z., Graafland, M. & Schijven, M. P. (2016), 'Systematic review on the effectiveness of augmented reality applications in medical training', *Surgical endoscopy* **30**(10), 4174–4183.

Baysal, S. & Duygulu, P. (2015), 'Sentioscope: a soccer player tracking system using model field particles', *IEEE Transactions on Circuits and Systems for Video Technology* **26**(7), 1350–1362.

Bellalouna, F. (2020), Industrial use cases for augmented reality application, *in* '2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)', IEEE, pp. 11–18.

Bernsen, N. O., Dybkjær, H. & Dybkjær, L. (1994), 'Wizard of oz prototyping: How and when', *Proc. CCI Working Papers Cognit. Sci./HCI, Roskilde, Denmark* .

Bielli, S. & Harris, C. G. (2015), A mobile augmented reality system to enhance live sporting events, *in* 'Proceedings of the 6th Augmented Human International Conference', ACM, pp. 141–144.

Billinghurst, M. (2021), 'Grand Challenges for Augmented Reality', *Frontiers in Virtual Reality* **2**, 12.
URL: *https://www.frontiersin.org/article/10.3389/frvir.2021.578080*

Billinghurst, M., Clark, A., Lee, G. & Others (2015), 'A survey of augmented reality', *Foundations and Trends®in Human–Computer Interaction* **8**(2-3), 73–272.

Billinghurst, M. & Duenser, A. (2012), 'Augmented reality in the classroom', *Computer* **45**(7), 56–63.

Billinghurst, M., Grasset, R. & Looser, J. (2005), 'Designing augmented reality interfaces', *ACM Siggraph Computer Graphics* **39**(1), 17–22.

Blauberger, P., Marzilger, R. & Lames, M. (2021), 'Validation of player and ball tracking with a local positioning system', *Sensors* **21**(4), 1465.

Boland, M. (2018), '990 Million ARKit & ARCore Smartphones in the Wild'.
URL: *https://arinsider.co/2018/12/27/990-million-arkit-arcore-smartphones-in-the-wild/*

Boland, M. (2021), 'Mobile AR Users Approach 800 Million'.
URL: *https://arinsider.co/2021/07/01/mobile-ar-users-approach-800-million/*

Bower, M., Howe, C., McCredie, N., Robinson, A. & Grover, D. (2014), 'Augmented Reality in education–cases, places and potentials', *Educational Media International* **51**(1), 1–15.

Bowman, D. A. & McMahan, R. P. (2007), 'Virtual reality: how much immersion is enough?', *Computer* **40**(7), 36–43.

Braud, T., Bijarbooneh, F. H., Chatzopoulos, D. & Hui, P. (2017), Future networking challenges: The case of mobile augmented reality, *in* '2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)', IEEE, pp. 1796–1807.

Bressa, N., Korsgaard, H., Tabard, A., Houben, S. & Vermeulen, J. (2021), 'What's the Situation with Situated Visualization? A Survey and Perspectives on Situatedness', *IEEE Transactions on Visualization and Computer Graphics* .

Brodlie, K. W., Carpenter, L. A., Earnshaw, R. A., Gallop, J. R., Hubbold, R. J., Mumford, A. M., Osland, C. D. & Quarendon, P. (2012), *Scientific visualization: techniques and applications*, Springer Science & Business Media.

Brooks, F. P. (1999), 'What's real about virtual reality?', *IEEE Computer graphics and applications* **19**(6), 16–27.

Buhl, H. U. & Winter, R. (2009), 'Full Virtualization–BISE's Contribution to a vision', *Business & Information Systems Engineering* **1**(2), 133–136.

Buric, M., Ivasic-Kos, M. & Pobar, M. (2019), Player tracking in sports videos, *in* '2019 IEEE International Conference on Cloud Computing Technology and Science (CloudCom)', IEEE, pp. 334–340.

Card, M. (1999), *Readings in information visualization: using vision to think*, Morgan Kaufmann.

Card, S. K. & Mackinlay, J. (1997), The structure of the information visualization design space, *in* 'Proceedings of VIZ'97: Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium', IEEE, pp. 92–99.

Carter, M., Allison, F., Downs, J. & Gibbs, M. (2015), Player identity dissonance and voice interaction in games, *in* 'Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play', pp. 265–269.

Cavallaro, R., Hybinette, M., White, M. & Balch, T. (2011), 'Augmenting Live Broadcast Sports with 3D Tracking Information', *IEEE MultiMedia Magazine* .
URL: *http://ieeexplore.ieee.org/document/6077863/?reload=true&arnumber=6077863*

Centieiro, P., Cardoso, B., Romão, T. & Dias, A. (2014), If you can feel it, you can share it: a system for sharing emotions during live sports broadcasts, *in* 'Proceedings of the 11th Conference on advances in computer entertainment technology', Vol. 2014- of *ACE '14*, ACM, pp. 1–8.

Centieiro, P., Romão, T. & Dias, A. E. (2015), Emotion Sharing during Live Sports Broadcasts: Studying Its Potential and the Users' Preferences, *in* 'Proceedings of the

12th International Conference on Advances in Computer Entertainment Technology', ACE '15, Association for Computing Machinery, New York, NY, USA.
URL: *https://doi.org/10.1145/2832932.2832986*

Chandler, T., Cordeil, M., Czauderna, T., Dwyer, T., Glowacki, J., Goncu, C., Klapperstueck, M., Klein, K., Marriott, K., Schreiber, F. & Others (2015), Immersive analytics, *in* '2015 Big Data Visual Analytics (BDVA)', IEEE, pp. 1–8.

Chapanis, A. (1991), 'To communicate the human factors message, you have to know what the message is and how to communicate it', *Human Factors Society Bulletin* **34**(11), 1–4.

Chatzopoulos, D., Bermejo, C., Huang, Z. & Hui, P. (2017), 'Mobile augmented reality survey: From where we are to where we go', *IEEE Access* **5**, 6917–6950.

Chen, C., Wang, O., Heinzle, S., Carr, P., Smolic, A. & Gross, M. (2013), Computational sports broadcasting: Automated director assistance for live sports, *in* '2013 IEEE International Conference on Multimedia and Expo (ICME)', IEEE, pp. 1–6.

Chen, K., Li, T., Kim, H.-S., Culler, D. E. & Katz, R. H. (2018), Marvel: Enabling mobile augmented reality with low energy and low latency, *in* 'Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems', pp. 292–304.

Chung, D. H. S., Legg, P. A., Parry, M. L., Bown, R., Griffiths, I. W., Laramee, R. S. & Chen, M. (2015), 'Glyph sorting: Interactive visualization for multi-dimensional data', *Information Visualization* **14**(1), 76–90.

Clay, V., König, P. & Koenig, S. (2019), 'Eye tracking in virtual reality', *Journal of eye movement research* **12**(1).

Cordeil, M., Cunningham, A., Bach, B., Hurter, C., Thomas, B. H., Marriott, K. & Dwyer, T. (2019), Iatk: An immersive analytics toolkit, *in* '2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)', IEEE, pp. 200–209.

Cordeil, M., Dwyer, T., Klein, K., Laha, B., Marriott, K. & Thomas, B. H. (2016), 'Immersive collaborative analysis of network connectivity: CAVE-style or head-mounted display?', *IEEE transactions on visualization and computer graphics* **23**(1), 441–450.

Cranmer, E. E., tom Dieck, M. C. & Fountoulaki, P. (2020), 'Exploring the value of augmented reality for tourism', *Tourism Management Perspectives* **35**, 100672.

Cutolo, F., Cattari, N., Fontana, U. & Ferrari, V. (2020), 'Optical see-through head-mounted displays with short focal distance: conditions for mitigating parallax-related registration error', *Frontiers in Robotics and AI* **7**.

da Silva, A. M., Gustavo, S. G. & de Medeiros, F. P. A. (2021), A Review on Augmented Reality applied to Sports, *in* '2021 16th Iberian Conference on Information Systems and Technologies (CISTI)', IEEE, pp. 1–6.

Danyluk, K., Ens, B., Jenny, B. & Willett, W. (2021), A design space exploration of worlds in miniature, *in* 'Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems', pp. 1–15.

Dasgupta, R., Dasgupta, R. & Srivastava (2018), *Voice User Interface Design*, Springer.

De Sá, M. & Churchill, E. (2012), Mobile augmented reality: exploring design and prototyping techniques, *in* 'Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services', pp. 221–230.

Deci, E. L. & Ryan, R. M. (2012), Handbook of Theories of Social Psychology: Volume 1, SAGE Publications Ltd, London.
URL: *https://sk.sagepub.com/reference/hdbk$_s$ocialpsychtheories*1

Desai, P. R., Desai, P. N., Ajmera, K. D. & Mehta, K. (2014), 'A review paper on oculus rift-a virtual reality headset', *arXiv preprint arXiv:1408.1173* .

Dewangan, O., Pinjarkar, L., Bonde, P. & Bagga, J. (2021), 'Advancement in Augmented and Virtual Reality', *Cognitive Behavior and Human Computer Interaction Based on Machine Learning Algorithm* pp. 211–239.

Dey, A., Billinghurst, M., Lindeman, R. W. & Swan, J. (2018), 'A systematic review of 10 years of augmented reality usability studies: 2005 to 2014', *Frontiers in Robotics and AI* **5**, 37.

Dünser, A., Grasset, R. & Billinghurst, M. (2008), *A survey of evaluation techniques used in augmented reality studies*, Human Interface Technology Laboratory New Zealand.

Dünser, A., Grasset, R., Seichter, H. & Billinghurst, M. (2007), 'Applying HCI principles to AR systems design', *Mixed Reality User Interfaces: Specification, Authoring, Adaptation (MRUI'07) Workshop Proceedings* pp. 37–42.

Ellis, G. & Dix, A. (2007), 'A taxonomy of clutter reduction for information visualisation', *IEEE transactions on visualization and computer graphics* **13**(6), 1216–1223.

ElSayed, N. A. M., Smith, R. T. & Thomas, B. H. (2016*a*), Horus eye: See the invisible bird and snake vision for augmented reality information visualization, *in* '2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)', IEEE, pp. 203–208.

ElSayed, N. A. M., Smith, R. T. & Thomas, B. H. (2016*b*), Situated analytics: Where to put the abstract data, *in* '28th Australian conference on human-computer interaction. https://doi. org/10.1016/j. jvlc', Vol. 6.

ElSayed, N., Thomas, B., Marriott, K., Piantadosi, J. & Smith, R. (2015), Situated analytics, *in* '2015 Big Data Visual Analytics', IEEE, pp. 1–8.

Ens, B. & Irani, P. (2016), 'Spatial Analytic Interfaces: Spatial User Interfaces for In Situ Visual Analytics', *IEEE computer graphics and applications* **37**(2), 66–79.

Evin, I., Pesola, T., Kaos, M. D., Takala, T. M. & Hämäläinen, P. (2020), 3pp-r: Enabling natural movement in 3rd person virtual reality, *in* 'Proceedings of the annual symposium on computer-human interaction in play', pp. 438–449.

Ferrer, C. D. R., Kitahara, I. & Kameda, Y. (2016), A Prospective Study About Enhancing Effect of VR in Soccer Training, Technical report, The Institute of Electronics, Information and Communication Engineers.

Fonseca, D., Martí, N., Redondo, E., Navarro, I. & Sánchez, A. (2014), 'Relationship between student profile, tool use, participation, and academic performance with the use of Augmented Reality technology for visualized architecture models', *Computers in human behavior* **31**, 434–445.

Fonseca, X., Lukosch, S. & Brazier, F. (2022), 'Design Framework for Social Interaction with Location-based Games', *International Journal of Serious Games* **9**(1), 59–81.

Ganapathy, S. (2013), Design guidelines for mobile augmented reality: User experience, *in* 'Human Factors in Augmented Reality Environments', Springer, pp. 165–180.

Gao, Q. H., Wan, T. R., Tang, W. & Chen, L. (2017), A stable and accurate markerless augmented reality registration method, *in* '2017 International Conference on Cyberworlds (CW)', IEEE, pp. 41–47.

Gimeno, J., Portalés, C., Coma, I., Fernandez, M. & Mart´, B. (2017), 'Combining traditional and indirect augmented reality for indoor crowded environments. A case study on the Casa Batlló museum', *Computers & graphics* **69**, 92–103.

Gorman, D. (2021), 'The state of play: 5 things to know about sports in 2021'.
URL: *https://blog.gwi.com/chart-of-the-week/5-things-sports/*

Greenwood, S., Perrin, A. & Duggan, M. (2016), 'Social media update 2016', *Pew Research Center* **11**(2), 1–18.

Grier, R. A. (2015), 'How high is high? a meta-analysis of nasa-tlx global workload scores', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* **59**(1), 1727–1731.
URL: *https://doi.org/10.1177/1541931215591373*

Grubert, J., Langlotz, T., Zollmann, S. & Regenbrecht, H. (2016), 'Towards Pervasive Augmented Reality: Context-Awareness in Augmented Reality', *IEEE Transactions*

on Visualization and Computer Graphics pp. 1–1.
URL: *http://ieeexplore.ieee.org/document/7435333/*

Gudmundsson, J. & Horton, M. (2017), 'Spatio-temporal analysis of team sports', *ACM Computing Surveys (CSUR)* **50**(2), 22.

Gui, J., Gu, D., Wang, S. & Hu, H. (2015), 'A review of visual inertial odometry from filtering and optimisation perspectives', *Advanced Robotics* **29**(20), 1289–1301.

Gul, S., Baker, L., Boult, R., Mills, S. & Zollmann, Z. (2021), Expert Sample Consensus Applied To Camera Localization for AR Sports Spectators, *in* 'Image and Vision Computing New Zealand (IVCNZ 2021)', IEEE, pp. 1–6.

Han, D.-I., tom Dieck, M. C. & Jung, T. (2018), 'User experience model for augmented reality applications in urban heritage tourism', *Journal of Heritage Tourism* **13**(1), 46–61.

Harrison, L., Reinecke, K. & Chang, R. (2015), Infographic aesthetics: Designing for the first impression, *in* 'Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems', ACM, pp. 1187–1190.

Hart, S. G. & Staveland, L. E. (1988), Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research, *in* 'Advances in psychology', Vol. 52, Elsevier, pp. 139–183.

Harter, A., Hopper, A., Steggles, P., Ward, A. & Webster, P. (2002), 'The anatomy of a context-aware application', *Wireless Networks* **8**(2), 187–197.

Hassan, H. G. (2016), Designing Infographics to support teaching complex science subject: A comparison between static and animated Infographics, PhD thesis, Iowa State University.

Henrysson, A., Marshall, J. & Billinghurst, M. (2007), Experiments in 3D interaction for mobile phone AR, *in* 'Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia', pp. 187–194.

Henze, N. & Boll, S. (2010), Evaluation of an off-screen visualization for magic lens and dynamic peephole interfaces, *in* 'Proceedings of the 12th international conference on Human computer interaction with mobile devices and services', pp. 191–194.

Heok, T. K. & Daman, D. (2004), A review on level of detail, *in* 'Proceedings. International Conference on Computer Graphics, Imaging and Visualization, 2004. CGIV 2004.', IEEE, pp. 70–75.

Herr, D., Reinhardt, J., Krüger, R., Reina, G. & Ertl, T. (2017), Immersive visual analytics for modular factory layout planning, *in* 'Workshop on Immersive Analytics of IEEE VIS'.

Hertzog, C., Sakurai, S., Hirota, K. & Nojima, T. (2020), Toward Augmented Reality Displays for Sports Spectators: A Preliminary Study, *in* 'Multidisciplinary Digital Publishing Institute Proceedings', number 1, p. 129.

Hettig, J., Engelhardt, S., Hansen, C. & Mistelbauer, G. (2018), 'AR in VR: assessing surgical augmented reality visualizations in a steerable virtual reality environment', *International journal of computer assisted radiology and surgery* **13**(11), 1717–1725.

Holloway, R. L. (1997), 'Registration error analysis for augmented reality', *Presence: Teleoperators & Virtual Environments* **6**(4), 413–432.

Huang, H., Wei Hong, L., Kher Hui, N., Brailsford, T. & O'Malley, C. (2018), Enhancing Reflective Learning Experiences in Museums Through Interactive Installations, *in* J. Kay & R. Luckin, eds, 'Rethinking learning in the digital age: Making the Learning Sciences count', pp. 776–783.

igroup presence questionnaire (IPQ) overview (n.d.).
URL: *http://www.igroup.org/projects/ipq/*

Ishii, H. (2008), 'The tangible user interface and its evolution', *Communications of the ACM* **51**(6), 32–36.

Iwata, T., Yamabe, T. & Nakajima, T. (2011), Augmented reality go: extending traditional game play with interactive self-learning support, *in* '2011 IEEE 17th International Conference on Embedded and Real-Time Computing Systems and Applications', Vol. 1, IEEE, pp. 105–114.

Izadi, S. (2019), 'Blending Realities with the ARCore Depth API'.
URL: *https://developers.googleblog.com/2019/12/blending-realities-with-arcore-depth-api.html*

Jacobs, M. C., Livingston, M. A. & State, A. (1997), Managing latency in complex augmented reality systems, *in* 'Proceedings of the 1997 symposium on Interactive 3D graphics', pp. 49—-ff.

Jang, S. H. & Hudson-Smith, A. (2012), Exploring mobile augmented reality navigation system for pedestrians, *in* 'Proceedings of the GIS Research UK 20th Annual Conference GISRUK'.

Jang, S., Woo, W. & Others (2005), 5W1H: Unified user-centric context, *in* 'The 7th International Conference on Ubiquitous Computing', Citeseer.

Jerald, J. (2004), 'Latency compensation for head-mounted virtual reality', *Department of Computer Science, University of North Carolina at Chapel Hill* .

Jingen Liang, L. & Elliot, S. (2021), 'A systematic review of augmented reality tourism research: What is now and what is next?', *Tourism and Hospitality Research* **21**(1), 15–30.

Jost, P., Cobb, S. & Hämmerle, I. (2020), 'Reality-based interaction affecting mental workload in virtual reality mental arithmetic training', *Behaviour & Information Technology* **39**(10), 1062–1078.

Kalkofen, D., Veas, E., Zollmann, S., Steinberger, M. & Schmalstieg, D. (2013), Adaptive Ghosted Views for Augmented Reality, *in* 'IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2013)'.

Kamphuis, C., Barsom, E., Schijven, M. & Christoph, N. (2014), 'Augmented reality in medical education?', *Perspectives on medical education* **3**(4), 300–311.

Kari, M., Grosse-Puppendahl, T., Coelho, L. F., Fender, A. R., Bethge, D., Schütte, R. & Holz, C. (2021), TransforMR: Pose-aware object substitution for composing alternate mixed realities, *in* '2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)', IEEE, pp. 69–79.

Keil, J., Schmitt, F., Engelke, T., Graf, H. & Olbrich, M. (2018), Augmented Reality Views: Discussing the Utility of Visual Elements by Mediation Means in Industrial AR from a Design Perspective, *in* 'International Conference on Virtual, Augmented and Mixed Reality', Springer, pp. 298–312.

Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J. & Melançon, G. (2008), Visual analytics: Definition, process, and challenges, *in* 'Information visualization', Springer, pp. 154–175.

Kim, K., Billinghurst, M., Bruder, G., Duh, H. B.-L. & Welch, G. F. (2018), 'Revisiting trends in augmented reality research: A review of the 2nd decade of ISMAR (2008–2017)', *IEEE transactions on visualization and computer graphics* **24**(11), 2947–2962.

Kim, S. L., Suk, H. J., Kang, J. H., Jung, J. M., Laine, T. H. & Westlin, J. (2014), Using Unity 3D to facilitate mobile augmented reality game development, *in* '2014 IEEE World Forum on Internet of Things (WF-IoT)', IEEE, pp. 21–26.

Kirkup, J. A., Rowlands, D. D. & Thiel, D. V. (2013), 'Indoor propagation investigation from a 2.4 GHz waist mounted beacon', *Procedia Engineering* **60**, 188–194.

Klein, G. & Murray, D. (2007), Parallel tracking and mapping for small AR workspaces, *in* '2007 6th IEEE and ACM international symposium on mixed and augmented reality', IEEE, pp. 225–234.

Klopfer, E. & Squire, K. (2008), 'Environmental Detectives—the development of an augmented reality platform for environmental simulations', *Educational technology research and development* **56**(2), 203–228.

Koba, M. (2013), 'Keeping fans in the stands is getting harder to do'.
URL: *https://www.cnbc.com/id/100886843*

Kohn, V. & Harborth, D. (2018), Augmented reality-A game changing technology for manufacturing processes?, *in* 'ECIS', p. 111.

Koskinen, I. & Frens, J. (2017), 'Research prototypes', *Archives of Design Research* **30**(3), 5–15.

Kourouthanassis, P. E., Boletsis, C. & Lekakos, G. (2015), 'Demystifying the design of mobile augmented reality applications', *Multimedia Tools and Applications* **74**(3), 1045–1066.

Kruijff, E., Swan, J. E. & Feiner, S. (2010), Perceptual issues in augmented reality revisited, *in* '2010 IEEE International Symposium on Mixed and Augmented Reality', IEEE, pp. 3–12.

Kunze, K., Minamizawa, K., Lukosch, S., Inami, M. & Rekimoto, J. (2017), 'Superhuman sports: Applying human augmentation to physical exercise', *IEEE Pervasive Computing* **16**(2), 14–17.

Lage, M., Ono, J. P., Cervone, D., Chiang, J., Dietrich, C. & Silva, C. T. (2016), 'StatCast Dashboard: Exploration of spatiotemporal baseball data', *IEEE computer graphics and applications* **36**(5), 28–37.

Langlotz, T., Nguyen, T., Schmalstieg, D. & Grasset, R. (2014), 'Next-Generation Augmented Reality Browsers: Rich, Seamless, and Adaptive', *Proceedings of the IEEE* **102**(2), 155–169.
URL: *http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6704316*

LaValle, S. M., Yershova, A., Katsev, M. & Antonov, M. (2014), Head tracking for the Oculus Rift, *in* '2014 IEEE international conference on robotics and automation (ICRA)', IEEE, pp. 187–194.

Lazard, A. & Atkinson, L. (2015), 'Putting environmental infographics center stage: The role of visuals at the elaboration likelihood model's critical point of persuasion', *Science Communication* **37**(1), 6–33.

Lee, C., Bonebrake, S., Bowman, D. A. & Höllerer, T. (2010), The role of latency in the validity of AR simulation, *in* '2010 IEEE Virtual Reality Conference (VR)', IEEE, pp. 11–18.

Legg, P. A., Chung, D. H. S., Parry, M. L., Jones, M. W., Long, R., Griffiths, I. W. & Chen, M. (2012), MatchPad: interactive glyph-based visualization for real-time sports performance analysis, *in* 'Computer Graphics Forum', Vol. 31, Wiley Online Library, pp. 1255–1264.

Li, L., Yu, F., Shi, D., Shi, J., Tian, Z., Yang, J., Wang, X. & Jiang, Q. (2017), 'Application of virtual reality technology in clinical medicine', *American journal of translational research* **9**(9), 3867.

Li, S., Blake, A. & Thomas, R. (2013), 'Modelling the economic impact of sports events: The case of the Beijing Olympics', *Economic Modelling* **30**, 235–244.

Liberati, N. (2018), 'Phenomenology, Pokémon Go, and other augmented reality games', *Human studies* **41**(2), 211–232.

Lim, Y.-k., Pangam, A., Periyasami, S. & Aneja, S. (2006), Comparative analysis of high-and low-fidelity prototypes for more valid usability evaluations of mobile devices, *in* 'Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles', pp. 291–300.

Lin, T., Yang, Y., Beyer, J. & Pfister, H. (2020), 'SportsXR–Immersive Analytics in Sports', *arXiv preprint arXiv:2004.08010* .

Lincoln, P., Blate, A., Singh, M., Whitted, T., State, A., Lastra, A. & Fuchs, H. (2016), 'From motion to photons in 80 microseconds: Towards minimal latency for virtual and augmented reality', *IEEE transactions on visualization and computer graphics* **22**(4), 1367–1376.

Lindlbauer, D., Feit, A. M. & Hilliges, O. (2019), Context-aware online adaptation of mixed reality interfaces, *in* 'Proceedings of the 32nd annual ACM symposium on user interface software and technology', pp. 147–160.

Liu, R., Zhang, J., Yin, K., Wu, J., Lin, R. & Chen, S. (2018), Instant SLAM initialization for outdoor omnidirectional augmented reality, *in* 'Proceedings of the 31st International Conference on Computer Animation and Social Agents', pp. 66–70.

Liu, S., Cui, W., Wu, Y. & Liu, M. (2014), 'A survey on information visualization: recent advances and challenges', *The Visual Computer* **30**(12), 1373–1393.

Liu, W., Li, X. & Huang, D. (2011), A survey on context awareness, *in* '2011 International Conference on Computer Science and Service System (CSSS)', IEEE, pp. 144–147.

Livingston, M. A. (2013), Issues in human factors evaluations of augmented reality systems, *in* 'Human factors in augmented reality environments', Springer, pp. 3–9.

Livingston, M. A. & Ai, Z. (2008), The effect of registration error on tracking distant augmented objects, *in* '2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality', IEEE, pp. 77–86.

Livingston, M. A., Swan, J. E., Gabbard, J. L., Hollerer, T. H., Hix, D., Julier, S. J., Baillot, Y. & Brown, D. (2003), Resolving multiple occluded layers in augmented reality, *in* 'The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.', IEEE, pp. 56–65.

Lopez-Caudana, E. O. (2011), 'Active Noise Cancellation: The Unwanted Signal and the Hybrid Solution', *Adaptive Filtering Applications, Dr. Lino Garcia (Ed.)* pp. 49–84.

Louis, T., Troccaz, J., Rochet-Capellan, A. & Bérard, F. (2019), Is it real? measuring the effect of resolution, latency, frame rate and jitter on the presence of virtual entities, *in* 'Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces', pp. 5–16.

Ludvigsen, M. & Veerasawmy, R. (2010), 'Designing Technology for Active Spectator Experiences at Sporting Events', *Proceedings of the 22Nd Conference of the*

Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction pp. 96–103.
URL: *http://doi.acm.org/10.1145/1952222.1952243*

Lungu, A. J., Swinkels, W., Claesen, L., Tu, P., Egger, J. & Chen, X. (2021), 'A review on the applications of virtual reality, augmented reality and mixed reality in surgical simulation: an extension to different kinds of surgery', *Expert review of medical devices* **18**(1), 47–62.

Mahmood, Z., Ali, T., Muhammad, N., Bibi, N., Shahzad, I. & Azmat, S. (2017), 'EAR: Enhanced Augmented Reality System for Sports Entertainment Applications.', *KSII Transactions on Internet & Information Systems* **11**(12).

Mann, S., Furness, T., Yuan, Y., Iorio, J. & Wang, Z. (2018), 'All reality: Virtual, augmented, mixed (x), mediated (x, y), and multimediated reality', *arXiv preprint arXiv:1804.08386* .

Masai, K., Sugiura, Y., Suzuki, K., Shimamura, S., Kunze, K., Ogata, M., Inami, M. & Sugimoto, M. (2015), AffectiveWear: towards recognizing affect in real life, *in* 'Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers', pp. 357–360.

Maybee, M. (2021), 'Satellite Events Could Be the Way of the Future: Here's Why'.
URL: *https://blog.pcnametag.com/satellite-events*

Mayer, R. E., Hegarty, M., Mayer, S. & Campbell, J. (2005), 'When static media promote active learning: Annotated illustrations versus narrated animations in multimedia instruction.', *Journal of Experimental Psychology: Applied* **11**(4), 256.

Melax, S., Keselman, L. & Orsten, S. (2013), Dynamics based 3D skeletal hand tracking, *in* 'Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games', p. 184.

Mercier-Ganady, J., Lotte, F., Loup-Escande, E., Marchal, M. & Lécuyer, A. (2014), The Mind-Mirror: See your brain in action in your head using EEG and augmented reality, *in* '2014 IEEE Virtual Reality (VR)', IEEE, pp. 33–38.

Milgram, P. & Kishino, F. (1994), 'A taxonomy of mixed reality visual displays', *IEICE TRANSACTIONS on Information and Systems* **77**(12), 1321–1329.

Mine, M. R. (1993), 'Characterization of end-to-end delays in head-mounted display systems', *The University of North Carolina at Chapel Hill, TR93-001* .

Mine, M. R., Van Baar, J., Grundhofer, A., Rose, D. & Yang, B. (2012), 'Projection-based augmented reality in disney theme parks', *Computer* **45**(7), 32–40.

Moniruzzaman, A. B. M. & Hossain, D. S. A. (2013), 'Comparative Study on Agile software development methodologies', *arXiv preprint arXiv:1307.3356* .

Mons, J. K. (2020), 'The top 10 basic features of future stadiums'.
URL: *https://sporttomorrow.com/the-top-10-features-of-future-stadiums/*

Moreland, K. (2012), 'A survey of visualization pipelines', *IEEE Transactions on Visualization and Computer Graphics* **19**(3), 367–378.

Morimatsu, S. (2019), Spectators' Experience Global Survey: Football Edition, Technical report, Deloitte Tohmatsu Consulting LLC, Tokyo, Japan.
URL: *https://www2.deloitte.com/sg/en/pages/operations/articles/football-spectator-experience-report.html*

Mulloni, A., Grubert, J., Seichter, H., Langlotz, T., Grasset, R., Reitmayr, G. & Schmalstieg, D. (2012), Experiences with the impact of tracking technology in mobile augmented reality evaluations, *in* 'MobileHCI 2012 Workshop MobiVis', Vol. 2, Citeseer.

Munzner, T. (2014), 'Marks and channels', *Visualization Analysis and Design* pp. 94–114.

Narasimhan, P., Drolia, U., Tan, J., Mickulicz, N. D. & Gandhi, R. (2015), 'The next-generation in-stadium experience (keynote)', *ACM SIGPLAN Notices* **51**(3), 1–10.

Narzt, W., Pomberger, G., Ferscha, A., Kolb, D., Müller, R., Wieghardt, J., Hörtner, H. & Lindinger, C. (2006), 'Augmented reality navigation systems', *Universal Access in the Information Society* **4**(3), 177–187.

Nguyen, H. (2017), '59% of sports fans would rather watch games from home'.
URL: *https://today.yougov.com/topics/sports/articles-reports/2017/08/08/59-sports-fans-would-rather-watch-games-home*

NintendoSoup (2019), 'POKEMON GO OFFICIALLY HITS 1 BILLION DOWNLOADS WORLDWIDE'.
URL: *https://nintendosoup.com/pokemon-go-officially-hits-1-billion-downloads-worldwide/*

Nuernberger, B., Ofek, E., Benko, H. & Wilson, A. D. (2016), SnapToReality: Aligning Augmented Reality to the Real World, *in* 'Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems', CHI '16, ACM, New York, NY, USA, pp. 1233–1244.
URL: *http://doi.acm.org/10.1145/2858036.2858250*

NuhoKibar, P. & Akkoyunlu, B. (2017), 'Fostering and assessing infographic design for learning: the development of infographic design criteria', *Journal of Visual Literacy* **36**(1), 20–40.

Olano, M., Cohen, J., Mine, M. & Bishop, G. (1995), Combatting rendering latency, *in* 'Proceedings of the 1995 symposium on Interactive 3D graphics', pp. 19—-ff.

Olsson, T. (2013), Concepts and subjective measures for evaluating user experience of mobile augmented reality services, *in* 'Human factors in augmented reality environments', Springer, pp. 203–232.

Olsson, T. & Salo, M. (2011), Online user survey on current mobile augmented reality applications, *in* '2011 10th IEEE International Symposium on Mixed and Augmented Reality', IEEE, pp. 75–84.

Ortman, E. & Swedlund, K. (2012), Guidelines for user interactions in mobile augmented reality, Master's thesis.

Osti, F., Santi, G. M. & Caligiana, G. (2019), 'Real Time Shadow Mapping for Augmented Reality Photorealistic Rendering', *Applied Sciences* **9**(11), 2225.

Otten, J. J., Cheng, K. & Drewnowski, A. (2015), 'Infographics and public policy: using data visualization to convey complex information', *Health Affairs* **34**(11), 1901–1907.

Owens, N., Harris, C. & Stennett, C. (2003), Hawk-eye tennis system, *in* '2003 International Conference on Visual Information Engineering VIE 2003', IET, pp. 182–185.

Palmas, F. & Klinker, G. (2020), Defining Extended Reality Training: A Long-Term Definition for All Industries, *in* '2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)', IEEE, pp. 322–324.

Papagiannis, H. (2017), *Augmented human: How technology is shaping the new reality*, " O'Reilly Media, Inc.".

Perin, C., Vuillemot, R. & Fekete, J.-D. (2013), 'SoccerStories: A kick-off for visual soccer analysis', *IEEE transactions on visualization and computer graphics* **19**(12), 2506–2515.

Perin, C., Vuillemot, R., Stolper, C. D., Stasko, J. T., Wood, J. & Carpendale, S. (2018), 'State of the Art of Sports Data Visualization', *Computer Graphics Forum* **37**(3), 663–686.

Petri, K., Bandow, N. & Witte, K. (2018), 'Using several types of virtual characters in sports-a literature survey', *International Journal of Computer Science in Sport* **17**(1), 1–48.

Poppe, R., Rienks, R. & van Dijk, B. (2007), Evaluating the future of HCI: challenges for the evaluation of emerging applications, *in* 'Artifical Intelligence for Human Computing', Springer, pp. 234–250.

Radianti, J., Majchrzak, T. A., Fromm, J. & Wohlgenannt, I. (2020), 'A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda', *Computers & Education* **147**, 103778.

Radu, I. (2012), Why should my students use AR? A comparative review of the educational impacts of augmented-reality, *in* '2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)', IEEE, pp. 313–314.

Ragan, E., Wilkes, C., Bowman, D. A. & Hollerer, T. (2009), Simulation of augmented reality systems in purely virtual environments, *in* '2009 IEEE Virtual Reality Conference', IEEE, pp. 287–288.

Reitmayr, G., Langlotz, T., Wagner, D., Mulloni, A., Schall, G., Schmalstieg, D. & Pan, Q. (2010), Simultaneous Localization and Mapping for Augmented Reality, *in* '2010 International Symposium on Ubiquitous Virtual Reality', pp. 5–8.
URL: *http://ieeexplore.ieee.org/document/5557942/*

Rematas, K., Kemelmacher-Shlizerman, I., Curless, B. & Seitz, S. (2018), Soccer on your tabletop, *in* 'Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition', pp. 4738–4747.

Ren, D., Goldschwendt, T., Chang, Y. & Höllerer, T. (2016), Evaluating wide-field-of-view augmented reality with mixed reality simulation, *in* '2016 IEEE Virtual Reality (VR)', IEEE, pp. 93–102.

Rengger, R. E. (1991), 'Human Aspects in Computing: Design and Use of Interactive Systems with Terminals, chapter Indicators of Usability based on performance'.

Renner, P. & Pfeiffer, T. (2017), [POSTER] Augmented Reality Assistance in the Central Field-of-View Outperforms Peripheral Displays for Order Picking: Results from a Virtual Reality Simulation Study, *in* '2017 IEEE international symposium on mixed and augmented reality (ISMAR-Adjunct)', IEEE, pp. 176–181.

Ritsos, P. D., Ritsos, D. P. & Gougoulis, A. S. (2011), Standards for augmented reality: A user experience perspective, *in* 'International AR standards meeting', pp. 1–9.

Rogers, R., Strudler, K., Decker, A. & Grazulis, A. (2017), 'Can Augmented-Reality Technology Augment the Fan Experience?: A Model of Enjoyment for Sports Spectators', *Journal of Sports Media* **12**(2), 25–44.

Rohmer, K., Jendersie, J. & Grosch, T. (2017), 'Natural environment illumination: Coherent interactive augmented reality for mobile and non-mobile devices', *IEEE transactions on visualization and computer graphics* **23**(11), 2474–2484.

Rokhsaritalemi, S., Sadeghi-Niaraki, A. & Choi, S.-M. (2020), 'A review on mixed reality: Current trends, challenges and prospects', *Applied Sciences* **10**(2), 636.

Román-Ibáñez, V., Pujol-López, F. A., Mora-Mora, H., Pertegal-Felices, M. L. & Jimeno-Morenilla, A. (2018), 'A low-cost immersive virtual reality system for teaching robotic manipulators programming', *Sustainability* **10**(4), 1102.

Ryan, R. M. (1982), 'Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory.', *Journal of personality and social psychology* **43**(3), 450.

Ryskeldiev, B., Ochiai, Y., Kusano, K., Li, J., Saraiji, Y., Kunze, K., Billinghurst, M., Nanayakkara, S., Sugano, Y. & Honda, T. (2021), Immersive Inclusivity at CHI: Design and Creation of Inclusive User Interactions Through Immersive Media, *in* 'Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems', pp. 1–4.

Sanna, A. & Manuri, F. (2016), 'A survey on applications of augmented reality', *Advances in Computer Science: an International Journal* **5**(1), 18–27.

Schall, G., Schmalstieg, D. & Junghanns, S. (2010), 'VIDENTE-3D visualization of underground infrastructure using handheld augmented reality', *GeoHydroinformatics: Integrating GIS and Water Engineering* pp. 207–219.

Schall, G., Zollmann, S. & Reitmayr, G. (2013), 'Smart vidente: Advances in mobile augmented reality for interactive visualization of underground infrastructure', *Personal Ubiquitous Comput.* **17**(7), 1533–1549.
URL: *https://doi.org/10.1007/s00779-012-0599-x*

Schmalstieg, D., Fuhrmann, A., Hesina, G., Szalavári, Z., Encarnaçao, L. M., Gervautz, M. & Purgathofer, W. (2002), 'The studierstube augmented reality project', *Presence: Teleoperators & Virtual Environments* **11**(1), 33–54.

Scholz, J. & Smith, A. N. (2016), 'Augmented reality: Designing immersive experiences that maximize consumer engagement', *Business Horizons* **59**(2), 149–161.

Schrepp, M. (2019), 'User Experience Questionnaire Handbook-Version 8', *URL: https://www.ueq-online.org/Material/Handbook.pdf* .

Schrepp, M., Hinderks, A. & Thomaschewski, J. (2017), 'Construction of a Benchmark for the User Experience Questionnaire (UEQ).', *IJIMAI* **4**(4), 40–44.

Schwind, V., Reinhardt, J., Rzayev, R., Henze, N. & Wolf, K. (2018), Virtual reality on the go? a study on social acceptance of vr glasses, *in* 'Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct', pp. 111–118.

Shen, S., Xu, K., Sotiriadis, M. & Wang, Y. (2022), 'Exploring the factors influencing the adoption and usage of Augmented Reality and Virtual Reality applications in tourism education within the context of COVID-19 pandemic', *Journal of hospitality, leisure, sport & tourism education* p. 100373.

Shimoda, S., Hayashi, M. & Kanatsugu, Y. (1989), 'New chroma-key imagining technique with hi-vision background', *IEEE Transactions on broadcasting* **35**(4), 357–361.

Shishido, H., Okada, Y., Kameda, Y., Koido, M. & Kitahara, I. (2020), 'Method of Multiview Video Switching for Soccer Game Analysis in Large Scale Space', *ITE Transactions on Media Technology and Applications* **8**(2), 70–80.

Shneiderman, B. (1996), The eyes have it: A task by data type taxonomy for information visualizations, *in* 'Proceedings 1996 IEEE symposium on visual languages', IEEE, pp. 336–343.

Shneiderman, B., Publ, A. & Norwood, N. J. (1993), '1.1 Direct manipulation: a step beyond programming', *Sparks of innovation in human-computer interaction* p. 17.

Sicat, R., Li, J., Choi, J., Cordeil, M., Jeong, W.-K., Bach, B. & Pfister, H. (2019), 'Dxr: A toolkit for building immersive data visualizations', *IEEE transactions on visualization and computer graphics* **25**(1), 715–725.

Sielhorst, T., Sa, W., Khamene, A., Sauer, F. & Navab, N. (2007), Measurement of absolute latency for video see through augmented reality, *in* '2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality', IEEE, pp. 215–220.

Singh, R. P., Javaid, M., Kataria, R., Tyagi, M., Haleem, A. & Suman, R. (2020), 'Significant applications of virtual reality for COVID-19 pandemic', *Diabetes & Metabolic Syndrome: Clinical Research & Reviews* **14**(4), 661–664.

Siricharoen, W. V. (2013), Infographics: the new communication tools in digital age, *in* 'The international conference on e-technologies and business on the web (ebw2013)', pp. 169–174.

Siricharoen, W. V. (2015), Infographic role in helping communication for promoting health and well-being, *in* 'Conference: proceedings of the second international conference on computer science, computer engineering, and education technologies (CSCEET2015). Kuala Lumpur, Malaysia'.

Skinner, P., Ventura, J. & Zollmann, S. (2018), Indirect Augmented Reality Browser for GIS Data, *in* '2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)', IEEE, pp. 145–150.

Slater, M. (2003), 'A note on presence terminology', *Presence connect* **3**(3), 1–5.

Smiciklas, M. (2012), *The power of infographics: Using pictures to communicate and connect with your audiences*, Que Publishing.

Soltani, P. & Morice, A. H. P. (2020), 'Augmented reality tools for sports education and training', *Computers & Education* **155**, 103923.

Spohrer, J. C. (1999), 'Information in places', *IBM Systems Journal* **38**(4), 602–628.

Stadium of the Future (2019), Technical report, Oracle Food and Beverage.
URL: *https://www.oracle.com/a/ocom/docs/dc/stadium-of-the-future-report-final.pdf*

Stauffert, J.-P., Niebling, F. & Latoschik, M. E. (2018), Effects of latency jitter on simulator sickness in a search task, *in* '2018 IEEE conference on virtual reality and 3D user interfaces (VR)', IEEE, pp. 121–127.

Stein, M., Janetzko, H., Lamprecht, A., Breitkreutz, T., Zimmermann, P., Goldlücke, B., Schreck, T., Andrienko, G., Grossniklaus, M. & Keim, D. A. (2018), 'Bring it to the pitch: Combining video and movement data to enhance team sport analysis', *IEEE transactions on visualization and computer graphics* **24**(1), 13–22.

Stropnik, V., Babuder, K., Črmelj, V., Pogačnik, M. & Others (2018), A Look into the Future of Sports: A Study of the Actual State of the Art-the Microsoft HoloLens and Augmented Reality, *in* '2018 International Conference on Broadband Communications for Next Generation Networks and Multimedia Applications (CoBCom)', IEEE, pp. 1–6.

Sun, X. & May, A. (2009), 'The role of spatial contextual factors in mobile personalization at large sports events', *Personal and Ubiquitous Computing* **13**(4), 293–302.

Sutherland, I. E. (1968), A head-mounted three dimensional display, *in* 'Proceedings of the December 9-11, 1968, fall joint computer conference, part I', ACM, pp. 757–764.

Sutton, J., Langlotz, T. & Plopski, A. (2022), 'Seeing Colours: Addressing Colour Vision Deficiency with Vision Augmentations using Computational Glasses', *ACM Trans. Comput.-Hum. Interact* **29**, 2022.
URL: *https://doi.org/10.1145/3486899*

Szalavári, Z. & Gervautz, M. (1997), The personal interaction Panel–a Two-Handed interface for augmented reality, *in* 'Computer graphics forum', Vol. 16, Wiley Online Library, pp. C335—-C346.

Talmaki, S. A., Dong, S. & Kamat, V. R. (2010), Geospatial databases and augmented reality visualization for improving safety in urban excavation operations, *in* 'Construction Research Congress 2010: Innovation for Reshaping Construction Practice', pp. 91–101.

Tatzgern, M. (2015), Situated Visualization in Augmented Reality, PhD thesis, Graz University of Technology.

Tatzgern, M., Kalkofen, D., Grasset, R. & Schmalstieg, D. (2014), Hedgehog labeling: View management techniques for external labels in 3D space, *in* '2014 IEEE Virtual Reality (VR)', IEEE, pp. 27–32.

Tatzgern, M., Orso, V., Kalkofen, D., Jacucci, G., Gamberini, L. & Schmalstieg, D. (2016), Adaptive information density for augmented reality displays, *in* '2016 IEEE Virtual Reality (VR)', IEEE, pp. 83–92.

Tavani, S., Billi, A., Corradetti, A., Mercuri, M., Bosman, A., Cuffaro, M., Seers, T. & Carminati, E. (2022), 'Smartphone assisted fieldwork: Towards the digital transition of geoscience fieldwork using LiDAR-equipped iPhones', *Earth-Science Reviews* p. 103969.

Teather, R. J., Pavlovych, A., Stuerzlinger, W. & MacKenzie, I. S. (2009), Effects of tracking technology, latency, and spatial jitter on object movement, *in* '2009 IEEE symposium on 3D user interfaces', IEEE, pp. 43–50.

Temoche, P., Ramirez, E. & Rodriguez, O. (2012), A low-cost data glove for virtual reality, *in* 'Proceedings of XI International Congress of Numerical Methods in Enginnering and Applied Sciences (CIMENICS).[internet]', Vol. 2012.

Thomas, B. H., Welch, G. F., Dragicevic, P., Elmqvist, N., Irani, P., Jansen, Y., Schmalstieg, D., Tabard, A., ElSayed, N. A. M., Smith, R. T. & Others (2018), Situated Analytics., *in* 'Immersive analytics', Vol. 11190, pp. 185–220.

Thompson, A. & Potter, L. E. (2017), Proposing augmentation of live sporting events with gamification and social sharing, *in* 'Proceedings of the 29th Australian Conference on Computer-Human Interaction', pp. 617–621.

Tiefenbacher, P., Lehment, N. H. & Rigoll, G. (2014), Augmented reality evaluation: A concept utilizing virtual reality, *in* 'International Conference on Virtual, Augmented and Mixed Reality', Springer, pp. 226–236.

Tsai, C.-H. & Huang, J.-Y. (2018), 'Augmented reality display based on user behavior', *Computer Standards & Interfaces* **55**, 171–181.

Tulshan, A. S. & Dhage, S. N. (2018), Survey on virtual assistant: Google assistant, siri, cortana, alexa, *in* 'International symposium on signal processing and intelligent recognition systems', Springer, pp. 190–201.

Tumler, J., Doil, F., Mecke, R., Paul, G., Schenk, M., Pfister, E. A., Huckauf, A., Bockelmann, I. & Roggentin, A. (2008), Mobile augmented reality in industrial applications: Approaches for solution of user-related issues, *in* '2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality', IEEE, pp. 87–90.

Ventura, R. B. & Porfiri, M. (2020), Galvanic Skin Response As a Measure of Engagement During Play in Virtual Reality, *in* 'Dynamic Systems and Control Conference', Vol. 84270, American Society of Mechanical Engineers, p. V001T17A003.

Vermeeren, A. P. O. S., Law, E. L.-C., Roto, V., Obrist, M., Hoonhout, J. & Väänänen-Vainio-Mattila, K. (2010), User experience evaluation methods: current state and development needs, *in* 'Proceedings of the 6th Nordic conference on human-computer interaction: Extending boundaries', pp. 521–530.

Vi, S., da Silva, T. S. & Maurer, F. (2019), User experience guidelines for designing hmd extended reality applications, *in* 'IFIP Conference on Human-Computer Interaction', Springer, pp. 319–341.

Ward, M. O. (2002), 'A taxonomy of glyph placement strategies for multidimensional data visualization', *Information Visualization* **1**(3-4), 194–210.

Weber, W. (2017), 'Interactive information graphics', *Information design: Research and practice* p. 172.

Weiser, M. (1998), The invisible interface: Increasing the power of the environment through calm technology, *in* 'International Workshop on Cooperative Buildings', Springer, p. 1.

Wensveen, S. & Matthews, B. (2014), Prototypes and prototyping in design research, *in* 'The routledge companion to design research', Routledge, pp. 262–276.

White, S. & Feiner, S. K. (2009*a*), SiteLens : Situated Visualization Techniques for Urban Site Visits, *in* 'Proceedings of the 27th international conference on Human Factors in computing systems (CHI2009)', ACM, pp. 1117–1120.
URL: *http://www.cs.columbia.edu/ swhite/pubs/white-2009-chi.pdf*

White, S. M. & Feiner, S. (2009*b*), *Interaction and presentation techniques for situated visualization*, Columbia University.

Willett, W., Jansen, Y. & Dragicevic, P. (2017), 'Embedded data representations', *IEEE transactions on visualization and computer graphics* **23**(1), 461–470.

Wither, J., Tsai, Y.-T. & Azuma, R. (2011), 'Indirect augmented reality', *Computers & Graphics* **35**(4), 810–822.

Wloka, M. M. (1995), 'Lag in multiprocessor virtual reality', *Presence: Teleoperators & Virtual Environments* **4**(1), 50–63.

Wohlgenannt, I., Simons, A. & Stieglitz, S. (2020), 'Virtual reality', *Business & Information Systems Engineering* **62**(5), 455–461.

Wong, A. Y.-Y., Ling, S. K.-K., Louie, L. H.-T., Law, G. Y.-K., So, R. C.-H., Lee, D. C.-W., Yau, F. C.-F. & Yung, P. S.-H. (2020), 'Impact of the COVID-19 pandemic on sports and exercise', *Asia-Pacific journal of sports medicine, arthroscopy, rehabilitation and technology* **22**, 39–44.

Wu, C.-W., Shieh, M.-D., Lien, J.-J. J., Yang, J.-F., Chu, W.-T., Huang, T.-H., Hsieh, H.-C., Chiu, H.-T., Tu, K.-C., Chen, Y.-T. & Others (2022), 'Enhancing Fan Engagement in a 5G Stadium With AI-Based Technologies and Live Streaming', *IEEE Systems Journal* .

Xi, N., Chen, J., Gama, F., Riar, M. & Hamari, J. (2022), 'The challenges of entering the metaverse: An experiment on the effect of extended reality on workload', *Information Systems Frontiers* pp. 1–22.

Yang, L. I., Huang, J., Feng, T., Hong-An, W. & Guo-Zhong, D. A. I. (2019), 'Gesture interaction in virtual reality', *Virtual Reality & Intelligent Hardware* **1**(1), 84–112.

Young, J. (2022), 'Sports teams are investing $10 billion in stadiums by 2030 — here's how they'll be different'.
URL: *https://www.cnbc.com/2022/01/08/the-future-of-sports-stadiums-smaller-arenas-more-tech-greener.html*

Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C.-L. & Grundmann, M. (2020), 'Mediapipe hands: On-device real-time hand tracking', *arXiv preprint arXiv:2006.10214* .

Zhang, R., Wu, L., Yang, Y., Wu, W., Chen, Y. & Xu, M. (2020), 'Multi-camera multi-player tracking with deep player identification in sports video', *Pattern Recognition* **102**, 107260.

Zheng, Z., Wang, B., Wang, Y., Yang, S., Dong, Z., Yi, T., Choi, C., Chang, E. J. & Chang, E. Y. (2017), Aristo: An augmented reality platform for immersion and interactivity, *in* 'Proceedings of the 25th ACM international conference on Multimedia', ACM, pp. 690–698.

Zollmann, S., Grasset, R., Langlotz, T., Lo, W. H., Mori, S. & Regenbrecht, H. (2020), 'Visualization Techniques in Augmented Reality: A Taxonomy, Methods and Patterns', *IEEE Transactions on Visualization and Computer Graphics* .

Zollmann, S., Langlotz, T., Loos, M., Lo, W. H. & Baker, L. (2019), Arspectator: Exploring augmented reality for sport events, *in* 'SIGGRAPH Asia 2019 Technical Briefs', pp. 75–78.

Zollmann, S., Langlotz, T., Regenbrecht, H., Button, C., Lo, W. H. & Mills, S. (2022), Augmented Reality for Sports Spectating and Coaching, *in* 'Interactive Sports Technologies', Routledge, pp. 96–111.

Zollmann, S. & Reitmayr, G. (2012), Dense depth maps from sparse models and image coherence for augmented reality, *in* 'Proceedings of the 18th ACM symposium on Virtual reality software and technology', ACM, pp. 53–60.

Appendix A

User study Documents

This section contains all user study document questionnaires, except for standardized questionnaires. Standardized questionnaires include the NASA Task Load Index (TLX) questionnaire, the igroup Presence Questionnaire (IPQ) and the User Experience Questionnaire (UEQ). Questionnaires that are repeated for different conditions are also omitted.

PARTICIPANT:_____ DATASET/ROTATION:_____ ORDER:_____ DATE:_________

Assumption:
The sides of the team are consistent throughout all task.
Orientation of the standard infographics might vary.

Task 1-1:
The dots you see on the infographics represents the rucks that occur during a game. Which area did most of the ruck events occur? Mark the region on the image below. Regions are separated by the white lines as seen in the image below.

Task 2-1:
Where is the orange team's "Player 14" initial position? Mark the position with an 'x' in the image below.

PARTICIPANT:_____ DATASET/ROTATION:_____ ORDER:_____ DATE:_________

Task 3-1:
Assuming the sides of the team are the same as previous task (orange on left and blue on right), Identify which side the top voted player is from and mark the correct region in the image below.

PARTICIPANT:_____ DATE:_________

For each of the questions below, tick the response that best characterises how you feel about the statement.

The following questions were adapted from the Unified Theory of Acceptance and Use of Technology [1].

		Strongly Disagree						Strongly Agree
Performance Expectancy (Usefulness)		**1**	**2**	**3**	**4**	**5**	**6**	**7**
1.	I would find situated infographics useful in a game.							
2	Using situated infographics allows me to find information I need more quickly.							
3.	Using the system increases my knowledge towards the game.							
Attitude toward using technology (Enjoyment)								
1.	Using situated infographics is a good idea.							
2.	Situated infographics makes the game more interesting.							
3.	I like using situated infographics							
Behavioral intentions to use the system (Preference)								
1.	I intend to use situated infographics in the next few games							
2.	I predict I would use situated infographics in the next few games							
3.	I plan to use the system in the next few games							

PARTICIPANT:_____ DATE:_________

Lastly, please fill in the remainder general questions.

1. Please rate your preferences on the standard 2D infographic vs the situated 3D infographics.

Standard 2D								Situated 3D

2. Please rate the usability of the standard 2D infographics.

Low Usability								High Usability

3. Please rate the usability of the situated 3D infographics.

Low Usability								High Usability

4. Please leave any feedback on the overall idea and concept.

5. Please leave any feedback on the implementation of the idea and concept.

6. Please leave any feedback on the experience of the user study.

THANK YOU!

Participant: Questionnaire 1

Questionnaire 1: Condition 1

There are no right or wrong answers, only your opinion counts.

1. How real did the AR experience seem to you?

   Completely real O O O O O O O Not at all

2. How much did your AR experience seem consistent with your real-world experience?

   Not consistent O O O O O O O Very consistent

3. I felt like I was just perceiving pictures.

   Fully disagree O O O O O O O Fully agree

4. I did not feel present in the augmented environment.

   Did not feel O O O O O O O Felt present

5. I still paid attention to the world outside the phone's view.

   Fully disagree O O O O O O O Fully agree

6. I felt present in the augmented environment.

   Fully disagree O O O O O O O Fully agree

7. I was not aware of my real environment.

   Fully disagree O O O O O O O Fully agree

Participant: Questionnaire 1

Questionnaire 1: Comparison between Condition 1 and Condition 2

1. Overall which condition did you prefer?
   ☐ Condition 1 ☐ Condition 2

2. Which of the two conditions did it feel like the virtual content was present in the real world?
   ☐ Condition 1 ☐ Condition 2

3. Which of the two conditions did you find more appealing?
   ☐ Condition 1 ☐ Condition 2

4. How significant is the difference you found between both conditions?

   Not significant Very significant
   ☐ ☐ ☐ ☐ ☐ ☐ ☐

5. Do you think the overall experience of the visualization feel different between both conditions?

   Not different Very different
   ☐ ☐ ☐ ☐ ☐ ☐ ☐

6. Does it matter to you which condition you were using to see the visualizations?

   Doesn't matter Matters a lot
   ☐ ☐ ☐ ☐ ☐ ☐ ☐

7. What are the main differences you perceived between the two conditions?

   ____
   ____
   ____

A few more questions…

Would you use the **Broadcast Style** and **AR visualization** application for your next game?

Broadcast	No								Yes
AR Vis	No								Yes

Do you often **feel lost** with what is going on in a game if you are just spectating on-site **without any technological assistance**?

No								Yes

Comparing to **broadcast styled** visualizations, did the **AR Visualizations** helped you keep track of where the action is going on?

No (Similar)								Yes (Helped a lot)

The current state of ARSpectator on mobile is designed to be used in short burst of time when additional information is required or when there is a break in the game. **How often** do you reckon you will use ARSpectator throughout an actual game?

Rarely (Once or twice)								Very Often (Every few minutes)

Do you prefer broadcast style visualization or AR visualization? Why?

☐ Broadcast Style ☐ AR Visualization

____

____

____

Do you have any feedback to improve what you experience overall? (Future visualizations, experience, ideas, interactions etc.)

____

____

____

Participant: Questionnaire 3: Overall

Please answer the following questions. Although the clips viewed are very short and might not be the team or sport of your choice, answer assuming that you are in a game watching a team of your choice playing.

Comparison of Methods Please rank the methods for each statement (1 = Best – 4 = Worst)		**On-site Viewing**	**On-site (Broadcast)**	**On-site (AR)**	**TV Broadcast**
1.	Overall game understanding (What is going on in the game, what is the current score, who scored etc.)				
2	Overall spatial understanding (Where is the action, which team is on which side etc.)				
3.	Overall team fanaticism (Do you feel more attached to your supporting team?)				
4.	Overall game satisfaction (Which method made the game more interesting)				
5.	Overall experience as a whole (Includes comfort, atmosphere, information, enjoyment, effort etc.)				

Please leave any comments or justifications regarding the methods above

____
____
____
____

Which method do you prefer to watch a sports game? Why?

On-site Viewing	**On-site (Broadcast)**	**On-site (AR)**	**TV Broadcast**

____
____
____
____

A few more questions…

Given the condition of the AR application with the factors at the noticeable level (The second time where you filled in the UEQ questionnaire), would you still use it in an actual game in the stadium? (Assuming the application is ready for use). Elaborate if needed.

☐ Yes ☐ No

____

____

____

Were you able to tell the differences between the 6 conditions you experience? Feel free to elaborate or give any additional justification.

☐ Yes, I can tell all 6 ☐ Yes, but only the 3 types ☐ No

____

____

____

Please write down anything else that is affecting your user experience but was not mentioned or if you have any ideas on what might potentially affect user experience.

____

____

____

A few more questions…

The visualizations that you experienced contain multiple levels of details (LOD), in which visualizations get more and more detailed with different LOD. An example of that is when visualizations appear with just graphical representation and then slowly reveal more details such as numbers and percentages. Do you think it is beneficial to have different LOD in each visualization as experienced? Please elaborate.

☐ Yes ☐ No

____
____
____
____
____

The following ratings are for the manual interface.

I thought there was too much inconsistency in this system

Strongly Disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ Strongly Agree

I find the interface and visualizations distracting

Strongly Disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ Strongly Agree

The following ratings are for the adaptive interface.

I thought there was too much inconsistency in this system

Strongly Disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ Strongly Agree

I find the interface and visualizations distracting

Strongly Disagree ☐ ☐ ☐ ☐ ☐ ☐ ☐ Strongly Agree

How “intelligent” do you think the system was at displaying the right visualization?

Very Poorly ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very Well

How “intelligent” do you think the system was at displaying visualizations in the right place?

Very Poorly ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very Well

How “intelligent” do you think the system was at displaying visualizations at the right time?

Very Poorly ☐ ☐ ☐ ☐ ☐ ☐ ☐ Very Well

Which interface would you prefer to use in a sports spectating scenario at the stadium? Please elaborate.

☐ Adaptive Interface (Automatic Visualizations) ☐ Manual Interface (Manual Trigger)

____

____

____

____

____

Please write down any other comments or feedback on using both types of interfaces.

____

____

____

____

____