Habilitation à diriger des recherches

Supporting Collaboration in Large Interactive Spaces

Favoriser la collaboration dans les grands espaces interactifs

Habilitation à diriger des recherches
de l'Université Paris-Saclay

présentée et soutenue à Gif-sur-Yvette, le 14 février 2024, par

Cédric FLEURY

Composition du jury

Anatole LÉCUYER Directeur de recherche, Inria Rennes, France	Rapporteur
Carman NEUSTAEDTER Professeur, Simon Fraser University, Canada	Rapporteur
Anthony STEED Professeur, University College London, Royaume-Uni	Rapporteur
Laurence NIGAY Professeur, Université Grenoble Alpes, France	Examinatrice
Michel BEAUDOUIN-LAFON Professeur, Université Paris-Saclay, France	Examinateur
Anastasia BEZERIANOS Professeur, Université Paris-Saclay, France	Présidente

Cédric Fleury, *Supporting Collaboration in Large Interactive Spaces,* Habilitation à Diriger des Recherches, Université Paris-Saclay © May, 2024

ACKNOWLEDGMENTS

First, I want to thank all the members of my Habilitation committee. I am extremely honored that they have accepted to be part of this committee. I would especially like to thank the three reviewers: Anatole Lécuyer, Carman Neustaedter and Anthony Steed for the time they spent reading the manuscript, and for their detailed and insightful reports. Special thanks to Carman for waking up so early to attend the defense because of the time difference. I would also like to thank the other members: Laurence Nigay for her challenging questions, Anastasia Bezerianos for accepting to chair the defense, and Michel Beaudouin-Lafon for mentoring me through the process, but also for his support and kindness during all these years.

Next, I want to deeply thank all my colleagues at Université Paris-Saclay, where I carried out most of the research that led to this Habiliation. Special thanks to Wendy Mackay for welcoming me as part of the InSitu and ExSitu teams, and for making these teams so lively and dynamic. I had the pleasure not only to share my office, but also to co-supervise a PhD student with Theophanis Tsandilas. I would like to thank Patrick Bourdot for inviting me to collaborate with the VENISE team and for co-supervising two PhD students with me. Michel, Wendy, Theophanis and Patrick definitely shaped the way I conducted my research. Thanks also to the past and current members of the InSitu, ExSitu and VENISE teams with whom I had the chance to work, and especially to Caroline Appert, Baptiste Caramiaux, Olivier Chapuis, Sarah Fdili Alaoui, Nicolas Ferey, Stéphane Huot, Huyen Nguyen, and Jeanne Vézien. I am extremely grateful to the engineers, Olivier Gladin, Nicolas Ladevèze and Alexandre Kabil, for their technical support on many projects and for helping me set up the technical system for my defense. As teaching was an important part of my work over these years, I would like to thank my colleagues at the Polytech Paris-Saclay engineering school for their dedicated involvement.

Then, I want to thank my new colleagues at IMT Atlantique with whom I have collaborated on research and teaching in the recent years. Special thanks to Thierry Duval, with whom I started my academic career, for his warm welcome at IMT Atlantique and his support. Thanks to Gilles Coppin, Mathieu Chollet and Cédric Dumas for co-supervising PhD students with me. I also thank Guillaume Moreau, Etienne Peillard and all the members of the INUIT team for our interesting discussions about research and many other topics.

Beyond these two institutions, I would like to thank my colleagues in France and abroad: Bruno Arnaldi and Valérie Gouranton with whom I started my career, and Tat-Jen Cham, Morten Fjeld, Henry Fuchs, Khanh-Duy Le, Tiberiu Popa, Bruce Thomas, and James Walsh with whom I collaborated on various projects.

Most importantly, I would like to warmly thank the PhD students I have had the pleasure of supervising over the years: Ignacio Avellino, Yujiro Okuya, Yiran Zhang, Arthur Fages, Thomas Rinnert and Aurélien Léchappé.

Finally, I want to thank my parents and my sister Delphine for their support all these years and for attending my defense. I would also like to thank my partner Delphine for her patience and unconditional support on a daily basis, and my daughter Clémence for distracting me from writing this manuscript.

CONTENTS

1

INTRODUCTION

The intensive use of digital technology leads to an exponential increase in the quantity, complexity and variety of data produced by our society. Many fields, such as science, industry, business or health, generate more and more data every year. For example, the telescope used for the Legacy Survey of Space and Time (LSST) will capture about a thousand 3.2-gigapixel images of the sky every night for ten years. The total amount of data collected over the ten years will reach about 60 petabytes of raw data [Obs22]. In healthcare, the amount of data collected each year is growing exponentially due to the accumulation of a wide variety of digital information including personal medical records, radiology images, clinical trial data and human genetics. New forms of data, such as 3D imaging, genomic sequences or biometric sensor recordings, further accentuate this trend. For instance, US healthcare data was estimated to be about 150 exabytes in 2011, but is now probably up to several zettabytes [RR14; Gui+16]. As a last example, the computer-aided design (CAD) model of a Boeing 777 is composed of 6 million parts and connectors, 350 million triangular faces to display, and 12 gigabytes of geometry data to store [XZ15].

This digital data provides unique opportunities to support scientific discovery, improve industrial processes, help decision-making, foster creation or encourage learning. However, enabling humans to handle and explore such a large quantity and variety of data is more than ever a major challenge. Although research work, especially in the domain of artificial intelligence (AI), focuses on the automatic processing of large amounts of such data, it is mandatory to keep humans involved in the analysis process. Users must keep control over how the data is processed and need to be able to understand the results provided by computers.

In this context, the need for computer-mediated collaboration has never been so high. On the one hand, processing or exploring complex data sets usually requires multiple collaborators to combine their expertise or share their knowledge. On the other hand, the current societal context, including pandemic situations and the green transition, requires groups of users to work remotely while intensively using digital tools. The COVID-19 pandemic has significantly increased our reliance on computer-mediated collaboration tools (+44%, according to Gartner [Rim22]), and some experts predict that this situation will largely persist after the pandemic [Blo22]. As a consequence, supporting collaboration among co-located and remote users when analyzing large amounts of data is also a key challenge.

My research investigates how large interactive spaces, such as wall-sized displays [And+11; Bea+12], immersive virtual reality systems [Cru+92] or augmented reality spaces [BK02], can foster collaboration on complex data sets. The ability of such systems to display large amounts of information, potentially in 3D, and to spatially organize that information offers new alternatives for interacting with digital content. These technologies start to be used in a wide range of application domains, such as scientific data analysis [**Fle+12**; PBC17; Kam+18], review of computer-aided design (CAD) models [Bou+10], monitoring of processes in control rooms [Sch+12], scheduling of complex events [Liu15] or pathology diagnosis [Rud+16]. They also provide large collaborative spaces for interaction and communication between multiple users, which can be useful for brainstorming and combining ideas during product design [**Oku+20**], crisis management [PBC18] or creative work [Str+99].

Nevertheless, large interactive spaces require us to rethink how users interact with computers and collaborate through them. In particular, they offer the opportunity to develop new forms of interaction and new solutions to foster collaboration by taking advantage of the large visualization space and physical space available to users. My research contributes to this evolution in four ways:

1. I propose new interaction paradigms to handle large displays and make use of the large physical space surrounding users. These paradigms contrast with traditional mouse-keyboard or touch-based interfaces by allowing multiple users to interact simultaneously, while moving freely within the system.
2. I explore collaborative interaction among multiple users within a shared interactive space. Such interaction provides users with dedicated features to enrich collaborative activities, while managing conflicts and interference that may arise when interacting in the same space or with the same content.
3. I create interactive systems that connect remote users across heterogeneous interactive spaces. These systems rely on specific technical solutions to synchronize complex data and transmit communication cues, including spatialized audio and 3D user representations.
4. I investigate video-mediated communication among remote collaborators in large interactive spaces. Such spaces involve specific constraints to deploy telepresence capabilities, but also offer unique possibilities to enhance remote users' perception and non-verbal communication.

1.1 TERMINOLOGY AND SCOPE

My research work is at the crossroads of Human-Computer Interaction (HCI), Virtual and Augmented Reality (VR/AR) and Computer-Supported Cooperative Work (CSCW). I study how humans communicate with computers, but also how humans communicate among them through computers. Although part of my research focuses on collaboration through immersive technologies including virtual and augmented reality, my contributions are broader in the HCI domain including work on interaction design and video-mediated communication.

Next, I define the key terminology used in this manuscript and hence clarify the scope of my research work:

Large interactive space. I choose this term to encompass both immersive and non-immersive systems that provide users with a vast physical space for interaction. I further define large interactive spaces and give examples in Section 2.1.

Mixed reality. Among the various terminologies that unify virtual and augmented reality, I opt to use the term "mixed reality" as originally defined by Milgram et al. [Mil+95]. The mixed reality continuum, recently revised by Skarbez et al. [SSW21], categorizes augmented reality (AR) and virtual reality (VR) systems according to the proportion of the real and virtual world perceived by users. I employ this term equivalently to "extended reality" in the manuscript.

Physical workspace. This workspace refers to the physical space where users interact within the system. It is typically defined by the available space in front of the displays or by the limits of the tracking system.

Virtual workspace. In mixed reality systems, the physical workspace is mapped onto a specific region of the virtual environment, referred to as the virtual workspace. This virtual workspace is the virtual counterpart of the physical workspace: users can travel everywhere in this virtual workspace by walking in the physical workspace.

Telepresence. I use the term "telepresence" to designate video-mediated communication systems that enhance users' feeling of being present in the same space. These systems usually rely on large or immersive displays and advanced sound synthesis. They contrast with conventional videoconferencing systems, using standard computers or mobile devices equipped with a single camera and a relatively small screen. Examples of such telepresence systems are presented in Section 2.3.1.

Collaborative virtual environment. I define collaborative virtual environments (CVEs) as distributed systems that enable remote users to meet in a shared mixed reality environment. This definition encompasses systems using either virtual reality or augmented reality technologies. It corresponds to what is sometimes referred to as social virtual reality (SVR). Section 2.3.2 describes collaborative virtual environments in more depth.

Awareness. When referring to awareness in this manuscript, I specifically focus on the awareness among collaborators. It includes all the perceptual cues that help users understand the position, actions and intentions of their collaborators, along with the information they aim to communicate.

User representation. I employ this terminology to refer to embodied representations that enable users to perceive remote collaborators in an interactive space. These representations can use either live video or 3D avatars to provide users with appropriate awareness of each other, depending on the context.

1.2 CONTEXT AND INSPIRATION

This manuscript presents my research activities since 2012. However, before that, my PhD work already focused on remote collaboration across immersive VR systems. While one part concentrated on distributed architecture for synchronizing virtual environments, another part explored collaborative interaction between users with heterogeneous devices. This work was part of a national project, named *Collaviz*, that provided valuable remote collaboration scenarios for scientific data analysis.

During my post-doctoral position (2012-2013), I explored 3D head reconstruction for telepresence at the *BeingThere Centre* [FSB14]. The ambition of this joint international laboratory was to connect remote rooms through stereoscopic wall-sized displays, which thus become glass windows between the rooms. Although my work focused on technological aspects, it enabled me to gain a better understanding of non-verbal communication, including facial expressions and eye gaze.

My position as an associate professor at Université Paris-Saclay (2013-2021) provided me with the opportunity to explore these research themes in various projects.

First, I collaborated on the Digiscope project[1], which created a network of ten interconnected platforms for interactive visualization of large datasets. This project was a unique occasion to design and test collaborative systems connecting remote platforms. It also provided various application domains, including scientific research, computer-aided design, decision support systems, and education. Second, I participated in projects involving engineers from the VR center of the PSA automotive company (now Stellantis[2]). It was a unique opportunity to access real collaborative design scenarios and interview engineers on their current practices.

In 2021, I joined IMT Atlantique as an associate professor. I continue to investigate collaboration in large interactive spaces. Part of this work is conducted in the context of Continuum[3], a follow up to the Digiscope project at a national level.

1.3 RESEARCH METHODOLOGY

The contributions described in this manuscript are supported by a wide range of empirical results. In particular, my work applied several HCI methods to investigate research questions and analyze findings. I learned this HCI methodology all along my research career, but I considerably improved my expertise during my position at Université Paris-Saclay. Beyond the traditional controlled experiments used to evaluate the proposed techniques or systems, I integrated user-centered methods involving potential users in the design of these techniques or systems whenever possible. These methods include participatory design, interviews and qualitative observations with low or high-fidelity prototypes. For example, we interviewed engineers from the PSA automotive company for the work on computer-aided design. We conducted qualitative observations and interviews with civil engineering students for the design of *ShapeCompare* (Section 3.1.2). We ran preliminary observations using low-fidelity prototypes for the design of *CamRay* (Section 4.2.1.2). When controlled experiments were not suitable to assess our solutions, we conducted user studies on more open-ended tasks to observe how users appropriate tools and compare their different strategies, as for evaluating *ARgus* (Section 4.2.3).

Although my training and expertise are more related to computer science, I have endeavored to ground my work with foundations in psychology and sociology. In particular, we designed collaborative systems based on the concept of *grounding in communication* proposed by Clark and Brennan [CB91]. This concept refers to the communication process required to build a *common ground* between users, including mutual knowledge, beliefs, and assumptions about the collaborative situations [CM81; CSB83]. Collaborative systems should either provide support to enhance the establishment of common ground, or compensate for communication cues that are lost in computer-mediated communication. In addition, our contributions on collaborative design build on specific previous work that studied collaborative design practices. We drew inspiration from the work of De Bono [De 67] on *lateral thinking*, and Stempfle and Badke-Schaub [SB02] on the thinking processes of design teams. Détienne's work [Dét06] also provided us with valuable insights on managing task interdependencies and multiple perspectives in design.

---

[1] http://www.digiscope.fr/en/

[2] https://www.stellantis.com/en/

[3] https://www.lri.fr/~mbl/CONTINUUM/en/

1.4 MANUSCRIPT OVERVIEW

This section describes the remaining chapters of the manuscript, and acknowledges the students and collaborators who contributed to this research work. After a brief chapter on related work, I divide my contributions into two parts: the first is related to interaction and co-located collaboration in large interactive spaces, while the second concerns remote collaboration across such spaces. For each part, I provide representative papers that illustrate my contributions to the related topic. These papers are attached at the end of the manuscript. Publications I have co-authored are highlighted in [**Bold**] in this manuscript.

Related work on large interactive spaces (Chapter 2). This chapter defines large interactive spaces and illustrates how users interact and collaborate in such systems. It also describes previous work that explores remote collaboration across large interactive spaces, including telepresence and collaborative virtual environments.

From interaction to collaboration in a shared interactive space (Chapter 3). Allowing each user to interact in large interactive spaces is a necessary step to support collaboration. This chapter introduces new interaction paradigms that provide users with the ability to master the unusual characteristics of these systems. Beyond individual interaction, it investigates how such systems can foster co-located collaboration by providing appropriate collaborative interaction among users.

The first section focuses on the large visualization space, studying how users can interact with 3D virtual objects on a wall-sized display and collaboratively explore a large number of these objects. The work on 3D interaction [**LF16**] was part of the master's thesis of J.-B. Louvet. The work on collaborative exploration [**Oku+20**] was part of the PhD thesis of Y. Okuya, co-supervised with P. Bourdot (CNRS senior researcher), and involved O. Gladin and N. Ladévèze (both engineers).

The second section concentrates on interaction in a 3D space, investigating 3D object deformation with haptic interaction and collaborative sketching in augmented reality. The first part on 3D object deformation [**Oku+18**a; **Oku+21**] was carried out during the PhD of Y. Okuya, co-supervised with P. Bourdot, in collaboration with N. Ladévèze. The second part on collaborative sketching [**FFT23**] was part of the PhD thesis of A. Fages, co-supervised with T. Tsandilas (Inria researcher).

The third section presents multiple navigation techniques that leverage the large physical space to maximize physical displacements and allow tangible interaction. It also investigates how these techniques can be extended to collaborative navigation involving two co-located users. All these contributions [**Zha+19**; **Zha+20**; **Zha+21**; **Zha+22**] were based on the PhD work of Y. Zhang, co-supervised with P. Bourdot, and involved S.-T. Ho (master's intern), T.T.H. Nguyen (associate professor at Université Paris-Saclay), and N. Ladévèze (engineer).

Representative papers:

- Interaction: *J.-B. Louvet, C. Fleury (2016). Combining Bimanual Interaction and Teleportation for 3D Manipulation on Multi-Touch Wall-sized Displays. Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST), 8 pages.*

- Collaboration: *Y. Okuya, O. Gladin, N. Ladévèze, C. Fleury, P. Bourdot (2020). Investigating Collaborative Exploration of Design Alternatives on a Wall-Sized Display. Proceedings of the ACM conference on Human Factors in Computing Systems (CHI), 12 pages.*
- Navigation: *Y. Zhang, N. Ladévèze, H. Nguyen, C. Fleury, P. Bourdot (2020). Virtual Navigation considering User Workspace: Automatic and Manual Positioning before Teleportation. Proc. of the ACM Symposium on Virtual Reality Software and Technology (VRST), 10 pages.*
- Collaborative navigation: *Y. Zhang, T.T.H. Nguyen, N. Ladévèze, C. Fleury, P. Bourdot (2022). Virtual Workspace Positioning Techniques during Teleportation for Co-located Collaboration in Virtual Reality using HMDs, Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR), 9 pages.*

Collaboration and awareness across remote spaces (Chapter 4). Remote collaboration is becoming crucial due to major changes in our society, including new work organization and the green transition. This chapter presents how large interactive spaces can foster collaboration among remote users. In particular, such collaboration requires technical solutions to connect heterogeneous platforms, as well as appropriate awareness and communication cues among the remote collaborators.

The first section focuses on the technical aspects of connecting remote users across heterogeneous systems. A distributed architecture synchronizing computer-aided design data [**Oku+18b**; **Oku+21**] was created during the PhD of Y. Okuya, co-supervised with P. Bourdot, in collaboration with N. Ladévèze and O. Gladin. As part of the DIGISCOPE project, we designed a system to explore 3D audio mappings between remote platforms [**Fyf+18**] with L. Fyfe (engineer), O. Gladin, and M. Beaudouin-Lafon (professor at Univ. Paris-Saclay). During my post-doc, I proposed a method to reconstruct remote users' head in 3D [**Fle+14**], supervised by H. Fuchs (professor at Univ. of North Carolina) and T.J. Cham (associate professor at NTU Singapore), and with the collaboration of T. Popa (post-doc at ETH Zurich).

The second section concentrates on video-mediated communication across large interactive spaces. The work on deictic gestures perception [**AFB15**] and telepresence across wall-sized displays [**Ave+17**] was part of the PhD work of I. Avellino, co-supervised with M. Beaudouin-Lafon, and involved W. Mackay (Inria senior researcher). The work about one-to-many telepresence [**Le+19**] was a collaboration with K.-D. Le (PhD student at Chalmers Univ. of Technology), I. Avellino, M. Fjeld (professor at Chalmers Univ. of Technology), and A. Kunz (professor at ETH Zurich). The work on multi-view collaboration with AR users [**FFT22b**] was part of the PhD thesis of A. Fages, co-supervised with T. Tsandilas.

Representative papers:

- Perceptual study: *I. Avellino, C. Fleury, M. Beaudouin-Lafon (2015). Accuracy of deictic gestures to support telepresence on wall-sized displays. Proceedings of the ACM conference on Human Factors in Computing Systems (CHI), 4 pages.*
- One-to-one telepresence: *I. Avellino, C. Fleury, W. Mackay, M. Beaudouin-Lafon (2017). CamRay: Camera Arrays Support Remote Collaboration on Wall- Sized Displays. Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI), 12 pages.*
- One-to-many telepresence: *K.-D. Le, I. Avellino, C. Fleury, M. Fjeld, A. Kunz (2019). GazeLens: Guiding Attention to Improve Gaze Interpretation in Hub-Satellite Collaboration. Proc. of the IFIP TC13 Conference on Human-Computer Interaction (INTERACT), 21 pages.*

- Telepresence with AR users: *A. Fages, C. Fleury, T. Tsandilas (2022). Understanding Multi-View Collaboration between Augmented Reality and Remote Desktop Users, Proc. of the ACM Conference on Computer-Supported Cooperative Work (CSCW), 27 pages.*

Future perspectives and closing remarks (Chapter 5). This chapter presents future directions for my research and concludes with some remarks regarding the future of collaboration in large interactive spaces.

2

RELATED WORK ON LARGE INTERACTIVE SPACES

The main objective of this chapter is to situate the context of my research work. As a consequence, it is not intended to be an exhaustive overview of the related work, but rather to provide examples that define and illustrate the foundations of my work. The first section proposes a definition of large interactive spaces and describes the different systems used in the work presented in this manuscript. The second section introduces some previous work that illustrates how users interact and collaborate in such systems. Finally, the last section presents related work addressing collaboration across remote interactive systems.

2.1 LARGE INTERACTIVE SPACES

Large interactive spaces include interactive systems that provide users with a large physical space, enabling them to move and interact in 3D. This large space can often accommodate multiple users, making these systems particularly well-suited to collaboration. By definition, they contrast with personal computers and mobile devices, which are generally designed for single users who remain static relative to the display. The key characteristics of large interactive spaces include:

- A **large visualization space**, which displays digital content in a large portion of the available 3D space, such as on a full wall or all around the users.
- A **large physical space**, which allows users to physically move around to explore the digital content and perform 3D interaction such as pointing, grabbing or manipulating virtual objects.
- Various **interaction devices**, which enable users to interact with the digital content from multiple locations in the 3D space. These devices range from touch devices to VR controllers.
- A **tracking system**, which detects the positions of both the users and the interaction devices in the 3D space.

With this definition, I decide to group immersive and non-immersive systems together, as they share many similarities in terms of interaction and collaboration. While hybrid systems do exist, non-immersive systems typically encompass 2D wall-sized displays and large digital tabletops, whereas immersive systems include all mixed reality devices. All these systems empower users to visualize and manipulate large volumes of complex data. They support physical navigation and 3D interaction to explore such data. Additionally, they can accommodate multiple users within the same interactive space and connect remote users.

In this section, I classify large interactive spaces into three categories that illustrate the wide range of systems available: (i) wall-sized displays, (ii) immersive virtual reality systems and (iii) augmented reality spaces. These three categories cover all the devices used in the research work described in this manuscript.

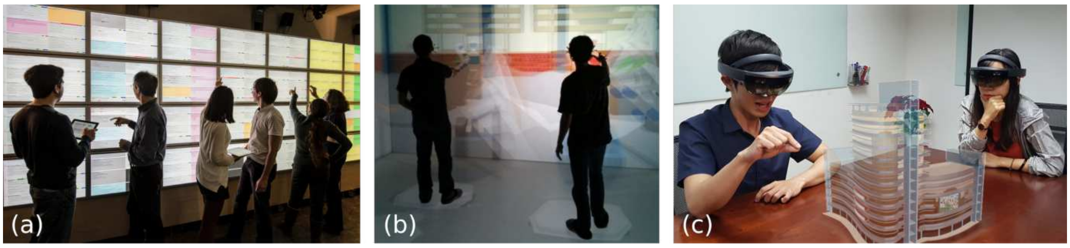


Figure 2.1: Large interactive spaces: (a) users preparing the CHI 2013 conference program on the WILD wall-sized display (© Inria), (b) users exploring a virtual factory in the EVE immersive system (© CNRS - VENISE), and (c) users reviewing a building 3D mockup in an augmented reality space (© Hoshinim - CC-BY-SA 4.0).

2.1.1 *Wall-sized displays*

Wall-sized displays typically consist of a large, ultra high-resolution display, as described by Andrews et al. [And+11] and Beaudouin-Lafon et al. [Bea+12]. They are often built using a large number of high-resolution screens, although a few use multiple projectors. Some previous studies have demonstrated their ability to improve performance in various tasks when compared to standard desktop computers. Czerwinski et al. [Cze+03] showed that larger screen space increases productivity and user satisfaction in everyday computer use, particularly when working with multiple windows. Bi and Balakrishnan [BB09] and Grudin [Gru01] corroborated these findings by showing that peripheral awareness facilitates tasks involving multiple windows. For sense-making tasks, Andrews et al. [AEN10] demonstrated that large screens enable users to employ a distributed cognitive process, which improves performance by allowing users to associate content meaning with spatial locations. Lischke et al. [Lis+15] proved the benefits of wall-sized displayed for information search tasks. Finally, other studies have shown that physical navigation provided by wall-sized displays improves spatial memory [JSH19] and user performance in visual search [BNB07] and data manipulation [Liu+14].

I used two wall-sized displays in my work: WILD (Figure 2.1-a) and WILDER. The WILD (Wall-sized Interaction with Large Datasets) platform was composed of an 8×4 grid of 30″ *Apple Cinema Displays* screens at the time we conducted the work presented in this manuscript. It measured $5.5m \times 1.8m$, with a resolution of 20480×6400 pixels. More recently, the WILD platform has been upgraded with $8K$ screens increasing the resolution to over 1 Gigapixel (61441×17240). It is controlled by a cluster of 16 computers, each managing two screens. A *VICON* infrared tracking system[1] can track the position and orientation of the users and interaction devices in front of the display.

The WILDER platform consists of a 15×5 grid of 21.6″ *Planar* screens[2] with $3mm$ ultra-thin bezels. It measures $5.9m \times 2m$, with a resolution of 14400×4800 pixels. It is controlled by a cluster of 10 computers, each managing a row of 8 or 7 screens. The platform also integrates a *VICON* tracking system, along with a *PQLabs* infrared frame[3] surrounding the display to provide multitouch capability.

---

[1] `https://www.vicon.com/`

[2] `https://www.planar.com/`

[3] `https://www.pqlabs.com/`

2.1.2 *Immersive virtual reality systems*

Virtual reality (VR) systems have been used for several decades to immerse users in virtual environments. They provide users with multi-sensory feedback, enabling them to perceive the virtual environment through multiple sensory cues, including visual, audio and haptic cues. This virtual environment is a digital world simulated by computers, which can present a large variety of information to the users. A more complete definition of virtual reality can be found in Fuchs et al. [FMG11]. When focusing on visual cues, the key characteristics of immersive VR systems include a wide field of regard, a wide field of view, high-resolution displays, and depth perception. The field of regard corresponds to the amount of the physical space surrounding users in which images are displayed, while the field of view refers to the viewing angle instantaneously perceived by users. Depth cues are essential for immersion, as they give users the feeling of being present in the 3D space of the virtual environment. In current VR systems, depth perception mainly relies on 3D stereoscopic vision and motion parallax. Motion parallax refers to the depth information conveyed by the relative movements of virtual objects in response to changes in the viewer's position. Providing users with the ability to move in the physical space of immersive systems is crucial for perceiving motion parallax.

A wide range of devices has been used to immerse users in virtual environments, including head-mounted displays, CAVE-like systems and various types of stereoscopic screens. The *CAVE* (Cave Automatic Virtual Environment), introduced in 1992 by Cruz-Neira et al. [Cru+92], was the first system using multiple projected screens surrounding users to increase the field of regard. Since then, the technology has improved considerably, and this type of platform has been deployed in many research laboratories and large companies, such as automotive and aerospace manufacturing companies. At the same time, head-mounted displays have also evolved dramatically, and a wide range of high-quality devices is now available to the general public.

In my research work, I used a CAVE-like system, named Eve (Figure 2.1-b), and head-mounted displays, including *HTC Vive* and *HTC Vive Pro Eye* headsets[4]. The Eve (Environnement Virtuel Evolutif) system is composed of four back-projected stereoscopic screens, measuring $4.8m \times 2.7m$ (front & floor) and $2.7m \times 2.7m$ (left & right). Each screen has a resolution of 1920×1080 pixels. The projectors are able to achieve both active stereovision and polarization image multiplexing, enabling two users to have their own stereoscopic view of the virtual environment. The system can thus support collaborative interaction between these two users, while providing both with correct motion parallax. Applications are executed on a server that distributes the rendering to four computers, each connected to a projector. An *ART* infrared tracking system[5] is used to capture the position and orientation of the users and interaction devices in the system. A *Scale-One* haptic device from Haption[6] is also available to interact with the virtual environment. This device consists of a *Virtuose* haptic arm[7] mounted on a 4-degree-of-freedom carrier, allowing users to interact anywhere in the physical space despite the limited range of the haptic arm.

---

[4] `https://www.vive.com/`

[5] `https://ar-tracking.com/`

[6] `https://www.haption.com/en/products-en/scale-one-en.html`

[7] `https://www.haption.com/en/products-en/virtuose-6d-en.html`

2.1.3 *Augmented reality spaces*

Augmented reality (AR) has the potential to transform any physical space into a large interactive space, as it can display virtual content anywhere in the physical space and allow users to interact with this content (Figure 2.1-c). Early work on augmented reality appeared at the beginning of the 1990s with systems, such as *KARMA* proposed by Feiner et al. [FMS93]. Early applications included industrial applications [CM92], medical applications [Fuc+96] or human-robot interaction [Mil+93]. Since then, AR hardware and software have drastically improved, as highlighted by Billinghurst et al. in their survey [BCL15]. Modern systems are now able to accurately track device positions, sense the physical world geometry in real time, and display content on lightweight wearable headsets. Virtual content can thus be seamlessly integrated into the real world with stable placement and convenient ways to interact with it. Moreover, augmented reality inherently supports collaboration between multiple users sharing the same physical space, since they can see each other in the real world. As a consequence, the number of AR applications has exploded, and they now reach a very wide range of domains, including marketing, medicine, education, entertainment, and architecture, as detailed in [BCL15].

With the large development of mixed reality headsets available to the general public, there is an increasing number of devices capable of delivering high-quality augmented reality experiences. These devices can be categorized into two types: optical and video see-through devices. Optical see-through devices enable users to see the real world through semitransparent screens, onto which the virtual content is overlaid. In contrast, video see-through devices display video feeds from cameras located on the device, and integrate the virtual content inside these video feeds. In the work presented in this manuscript, I used optical see-through devices, and more specifically the *Hololens 2* headsets from Microsoft[8].

2.2 INTERACTION IN LARGE INTERACTIVE SPACES

Large interactive spaces require specific techniques to interact with digital content due to their unique characteristics. In particular, they involve going beyond conventional 2D interfaces and keyboard-mouse interaction, as users must handle large visualization spaces while moving within the system. In this section, I present some previous work illustrating how users can interact with both 2D and 3D content in large interactive spaces. I also detail other related work that explores how multiple users can interact in the same interactive space.

2.2.1 *Interaction with 2D content*

In this subsection, I concentrate on how to interact with 2D digital content displayed on large screens, such as wall-sized displays or digital tabletops. A first solution is to interact at close proximity to the display through direct touch. However, this solution requires users to travel long distances along the screen and can lead to difficulties in accessing some areas, such as the very top or bottom of the screen. To solve this issue, Bezerianos and Balakrishnan [BB05a] introduced a dedicated

---

[8] https://www.microsoft.com/hololens/

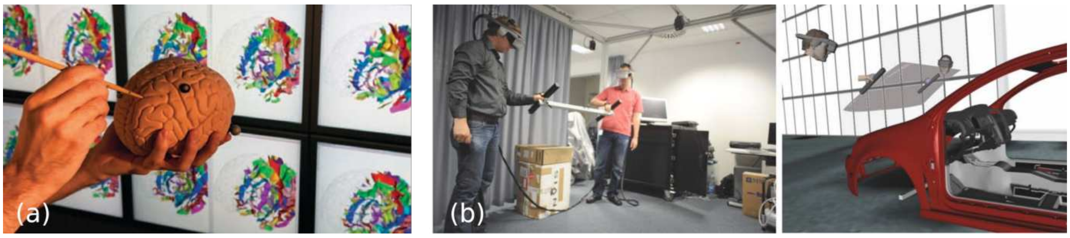


Figure 2.2: Tangible interaction in large interactive spaces: (a) a 3D brain model is used to control brain scans on a wall-sized display (Figure from Beaudouin-Lafon [Bea11]), and (b) a shared prop is used to manipulate a virtual car windshield by co-located users immersed in VR (Figure from Salzmann et al. [SJF09]).

widget on the display to attract far away items close to users. In a similar spirit, *Canvas portals* [BB05b] enable users to interact from distant parts of the display, using interactive portals that replicate the digital content of remote display areas.

To fully take advantage of wall-sized display characteristics, it is common to step away from the screen to get an overview of the displayed data. However, users still need to interact with the display content even when located at a distance. Previous work has explored interaction techniques to achieve mid-air pointing on wall-sized displays. Nancel et al. [Nan+15] provide a comprehensive overview of these techniques. Other studies have investigated the use of mobile devices to interact at a distance with wall-sized displays, ranging from smartwatches [Hor+18] to tablets [Kis+17] or tabletops [Bea11]. For example, *Smarties* [CBF14] proposes a generic solution for using touch on mobile devices to interact with wall-sized displays. It provides an interface on the mobile device, as well as a communication protocol for sending input to a wall-sized display to control interaction cursors or other specific elements. Lastly, 3D interaction with tangible props can also be used to interact with digital content on wall-sized displays. For instance, Beaudouin-Lafon [Bea11] described a scenario in which neuroscientists manipulate a stick in relation to a physical 3D brain model (Figure 2.2-a). The stick mimics a camera pointing at the brain and controls the orientation of brain scans distributed on the wall-sized display. *WallTokens* [CAC21] also proposes to use tangible objects that can be slid and attached to a wall-sized display to manipulate the digital content.

2.2.2 *Interaction with 3D content*

A large number of interaction techniques have been designed to interact with 3D content in mixed reality environments. An exhaustive list of 3D interaction techniques can be found in the book by LaViola et al. [LaV+17]. These techniques are usually classified in three categories, as proposed by Hand [Han97]: (i) selection and manipulation, (ii) navigation and (iii) system control. System control consists of sending commands to the application for requesting specific actions, changing the interaction mode, or modifying some parameters.

To select and manipulate virtual objects, we can use two approaches as for the interaction with 2D content. The first one involves traveling close to the virtual objects and performing direct manipulation. The simplest solution to achieve this is to use a virtual hand mimicking the movements of a user's real hand, as studied by Jacoby et al. [JFH94]. However, the travel actions required to reach the objects can be inconvenient and increase interaction complexity. The second approach thus takes advantage of being in a virtual environment to interact with objects at a distance with, for example, the *"Go-Go" technique* [Pou+96] or a virtual ray [Min95]. Argelaguet and Andujarb [AA13] presented an extensive survey of pointing techniques for 3D object selection in mixed reality.

A wide variety of navigation techniques have been proposed for traveling in virtual environments. An obvious approach is to let users walk, with their physical movements mapped to virtual displacements. This has many benefits [LaV+17], such as being easy to learn and use, increasing immersion as studied by Usoh et al. [Uso+99], reducing motion sickness, and promoting spatial understanding. However, as the physical space is limited, users can only cover short distances. Redirect walking, originally proposed by Razzaque et al. [RKW01], guides users away from the physical boundaries, enabling them to travel longer distances in the virtual environment. This technique either deforms the users' spatial perception, making them follow a circular path in the physical space [Ste+10] or reorients the virtual environment while users are looking at distractors [CF17]. Nevertheless, real walking is not always feasible or desired, especially in small physical spaces. In such cases, steering metaphors offer an alternative, allowing users to continuously control their direction and speed in the virtual environment. The direction and speed can be defined by a simple device such as a gamepad or based on users' body position [Che+13; KRF11]. Selection-based metaphors can also be used to choose a destination and reach it instantaneously, saving travel time. For example, teleportation techniques using a virtual ray to select the destination are now common in many VR headset applications, as they reduce motion sickness [Wei+18; Hab+18].

2.2.3 *Collaborative interaction*

Previous work has explored co-located collaboration in large interactive spaces. *Liveboard* [Elr+92] introduced a digital whiteboard with pen interaction and is probably one of the first vertical displays supporting group interaction. Jakobsen and Hornbæk [JH14] studied a collaborative problem-solving task on a wall-sized display and identified six distinct collaboration styles. They highlighted that physical proximity of participants is closely related to how tightly coupled they work. In later work [JH16], they assessed the impact of two input modalities on collaboration: touch input on the display and mouse interaction at a distance. They found that wall-sized displays can afford equal participation regardless of input modality. Although touch input seems well suited to collaboration allowing users to negotiate for space, it can lead to more interference and conflicts than with mouse input. Nevertheless, none of these studies provide interaction techniques designed to support collaborative activities. Liu et al. [Liu+16] studied different collaboration styles with pairs of participants in a data manipulation task on a wall-sized display. They varied the interaction and communication capabilities of each participant of the pair. They also included a technique that supports collaborative interaction by allowing participants to assist their partner with data manipulation. The results show the benefits of collaborative interaction, which enables participants to collaborate more tightly even when not in close proximity. Based on these findings, they proposed *CoReach* [Liu+17], a set of cooperative touch gestures that combine input from multiple users, allowing them to show or pass content to each other, as well as group multiple items on the wall-sized display. *GroupTogether* [MHG12] detects collaborators' spatial formations (*F-formations*) and how they orient and tilt mobile devices (*micro-mobility*). It uses this spatial information to facilitate cross-device interaction between collaborators' tablets and a wall-sized display. This is an interesting example of how sociological constructs can be leveraged to support appropriate collaboration techniques.

Other studies have concentrated on co-located collaboration in immersive VR systems. A few CAVE-like systems are able to display multiple stereoscopic views of the virtual environment, thus supporting multiple users. For example, the EVE system (Section 2.1.2) provides perspectively correct views for two users, while the *C1x6* system [Kul+11] can accommodate up to six users. This system also explored various navigation techniques to enable all six users to travel together in the virtual environment. Aguerreche et al. [ADL10b] proposed using a reconfigurable tangible device for collaborative manipulation in a CAVE-like system. Co-located collaboration can also be beneficial when multiple users equipped with VR head-mounted displays are in the same room. Although they cannot see each other, they can still hear each other and feel the force applied by others when sharing tangible objects. Such shared tangible objects can simulate users holding virtual objects together in the virtual environment, as studied by Salzmann et al. [SJF09] with a virtual car windshield (Figure 2.2-b). Navigation can be challenging in such co-located situations, as it is mandatory to maintain the spatial relationship between users while they travel the virtual environment. *Multi-Ray Jumping* [WKF19] provides a collaborative teleportation technique that preserves this spatial relationship.

Augmented reality inherently supports co-located collaboration as users can see each other, providing mutual awareness. As a consequence, early work on augmented reality already explored collaborative systems that let co-located users see and interact with shared AR content, such as *TransVision* [Rek96], *Shared Space* [BWF98], or *Studierstube* [Sza+98]. Later, Billinghurst et al. [Bil+02] developed a collaborative AR system including tangible elements that users can manipulate to interact with AR content. They showed that participants' behaviors with this system are closer to unmediated collaboration than with a large 2D display. Kiyokawa et al. [Kiy+02] studied communication behaviors of co-located users with various AR devices and AR content placement. They found that optical see-through headsets with a task space situated between participants may produce the most natural collaboration. This can be explained by the fact that participants could better perceive non-verbal communication cues with this combination of device and placement. A recent survey by Sereno et al. [Ser+22] compiles notable research work on co-located collaboration in AR over the last decade. While all these techniques enable users to act on shared AR content, very few of them support collaborative interaction or provide solutions to mitigate interference and conflicts during interaction. Oda and Feiner [OF09] introduced a redirected motion system designed to prevent interference between two users, but it is dedicated to hand-held AR devices.

2.3 COLLABORATION ACROSS REMOTE INTERACTIVE SPACES

A vast body of research in computer-supported cooperative work has explored remote collaboration. I distinguish two categories among previous work that studied collaboration across remote large interactive spaces. The first category concerns telepresence and video-mediated communication systems, which enable users to share video across remote locations. The second category involves collaborative virtual environments and aims to immerse remote users in a shared virtual environment using mixed reality technologies.

2.3.1 *Telepresence systems*

Video plays a crucial role in supporting non-verbal cues, turn-taking and shared understanding of the collaborative situation among remote users, as emphasized by Isaacs & Tang [IT93]. Similarly, Monk & Gale [MG02] observed that gaze awareness provides an alternative non-linguistic channel for checking mutual understanding among remote collaborators. When engaged in collaborative tasks, seeing each other's faces improves the negotiation of common ground, as shown by Veinott et al. [Vei+99]. Properly conveying gaze direction is challenging in such systems, due to the disparity between camera and viewer positions, as well as the fact that video is displayed on a flat screen. Previous work has explored solutions to support the correct interpretation of gaze direction. For instance, *Hydras* [SBA92] used multiple mobile devices combining a screen and a camera. These devices can be placed in a way that reflects remote collaborators' positions, thus preserving eye contact. *Multiview* [NC05] relies on a multi-view screen and multiple cameras to display an individual view with a correct perspective for each user. Pan and Steed [PS14] designed a cylindrical screen to preserve gaze direction by displaying perspective-correct images for multiple viewpoints around a conference table. Nevertheless, most of these systems consider static users sitting around a table.

Other work has focused on users moving in their physical space, using wall-sized displays to simulate a large glass window between the remote spaces. Willert et al. [Wil+10] connected two wall-sized displays by capturing video through a 2D grid of cameras and displaying video remotely with a motion parallax effect when the viewer moves. Consequently, this system supports only one user at each remote location. Dou et al. [Dou+12] created a similar setup by using a wide-angle camera to capture the background and multiple video+depth cameras to capture users in the foreground. Users are segmented using the depth data and overlaid on the wide-angle video, preserving eye contact with remote collaborators.

When collaborating remotely, users often need to work on shared digital content. Early work on telepresence investigated how to provide users with the same interaction space. *VideoDraw* [TM90] was a shared drawing tool that overlaid the

Figure 2.3: Telepresence across large interactive spaces: (a) *MirrorBlender* blends video feeds and shared screens from remote collaborators (Figure from Grønbæk et al. [Grø+21]), and (b) *t-Room* uses screens arranged in a circle to overlay remote collaborators' video feeds onto shared digital content (Figure from Luff et al. [Luf+15]).

drawing with a shadow of a remote collaborator's arm. It thus allows users to perceive the collaborator's hand gestures. It also includes a screen displaying the remote collaborator's face on top of the drawing area. *VideoWhiteboard* [TM91] improved on this system by overlaying the shared drawing with a shadow of the collaborator's full body. *Clearboard* [IK92] extended this idea by blending the shared drawing with the video of the remote collaborator using transparency. The system could thus convey facial expressions and gaze awareness. More recently, *MirrorBlender* [Grø+21] designed a video-conferencing system that enables users to reposition, resize and blend with transparency multiple video feeds and shared screens from remote collaborators (Figure 2.3-a). Finally, *t-Room* [Luf+15] is a telepresence system connecting two remote locations with large screens arranged in a circle around a digital tabletop (Figure 2.3-b). Cameras are attached on top of each screen and capture users inside the interactive space. Shared digital content on the screens is overlaid with the remote video in a way that preserves the physical relations between the users and digital objects. However, this system has the drawback of requiring exactly the same complex setup at both locations.

2.3.2 *Collaborative virtual environments*

Collaborative virtual environments (CVEs) enable remote users to share virtual content, whether they use AR or VR devices. Early work primarily focused on distributing and synchronizing the virtual environment data across remote locations, as studied during my PhD [**Fle+10c**; **Fle+10b**]. Practical solutions such as Ubiq [Fri+21] are now available for creating CVEs. Recent work mainly concentrates on improving awareness and collaborative interaction among remote collaborators.

In such environments, providing mutual awareness is mandatory to support effective collaboration and social presence among remote users. A large number of studies have investigated the use of avatars to provide embodied representations of users in the virtual environment. Current systems can now create highly realistic avatars with simple technical setups, such as those proposed by Bartl et al. [Bar+21]. However, other authors prefer using low realism or cartoon-like avatars [FM21; KMI19], as realistic avatars can induce an Uncanny Valley effect [MMK12]. There is not a clear consensus regarding the impact of these two approaches on social presence, and it could highly depend on the collaborative context, as highlighted by Yoon et al. [Yoo+19]. In addition, a few systems have used live video [Ben+95] or real-time 3D reconstruction [Bec+13] to represent users in virtual environments (Figure 2.4-a). Congdon et al. [Con+23] compared the effects of video and 3D avatar representations on user trust in a CVE. The results are not clear-cut, but it appears that animated 3D avatars can perform as well as full-body video and better than head-and-shoulder video. Once again, the collaborative context and the environment surrounding the users' representations are likely to have significant effects on these results, and further studies would be required.

While remote users can interact in CVEs using the individual interaction techniques presented in Section 2.2.2, additional techniques can be useful to support collaborative practices. In particular, it is crucial to manage conflicts that may arise when users manipulate the same virtual objects and to allow cooperative manipulation of these objects. To prevent conflicts, *Spacetime* [Xia+18] creates a parallel version of objects whenever a conflict occurs. Users can thus manipulate their own

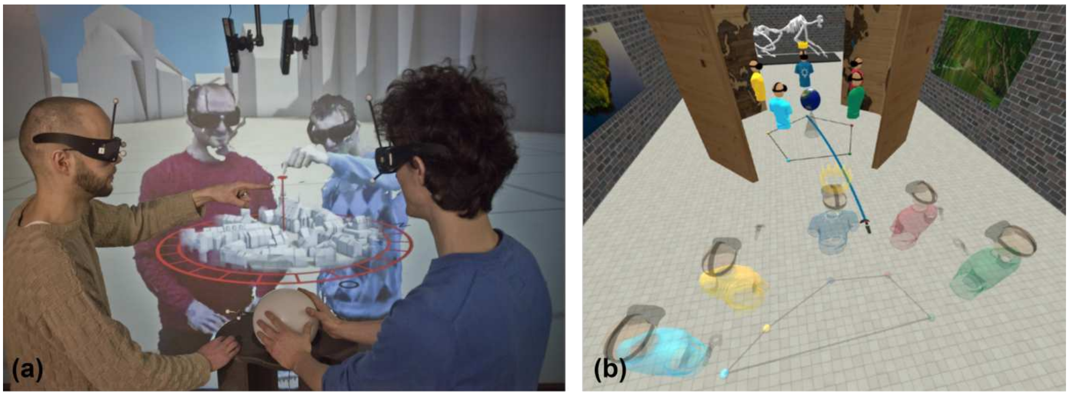


Figure 2.4: Collaborative interaction in virtual environments: (a) a World-In-Miniature is used to move collaborators in the environment (Figure from Beck et al. [Bec+13]), and (b) a specific teleportation technique enables a group to travel together while adjusting its spatial formation (Figure from Weissker et al. [WF21]).

version of an object and review all the created versions before choosing the final one. Pinho et al. [PBF02] propose to combine users' actions by separating the degrees of freedom they control. To enable simultaneously manipulation of the same degrees of freedom, Noma and Miyasato [NM97] used a physical simulation to compute the force applied by each user to the virtual object. *SkeweR* [DLT06] enables two users to simultaneously grab a virtual object using two control points. The position and orientation of the object are determined by the positions of these control points. Aguerreche et al. [ADL09] suggest adding a third control point to more precisely manipulate the object orientation along the three axes. One user thus manipulates two control points, while the other manipulates the third one. During my PhD, I used this 3-hand manipulation technique to assess the co-manipulation of a shared virtual object by two remote users in CAVE-like systems [**Fle+12**].

Collaboration navigation in CVEs can be challenging because it can be difficult for users to meet and travel together in the virtual environment. It is even more complex when teleportation is involved, as collaborators may disappear when they "jump" to another location in the environment. To overcome this issue, *Spacetime* [Xia+18] proposes to use parallel versions of the collaborators' avatars. When a collaborator teleports, a parallel avatar remains at the original location, allowing any user to select it and follow the collaborator. Additionally, *Spacetime* enables users to create a parallel avatar by grabbing a collaborator's avatar. Users can move this parallel avatar to a new location where they want to show something to the collaborator. The collaborator then receives a notification and can decide to travel automatically to this new location. A World-In-Miniature (WIM), originally proposed by Stoakley et al. [SCP95], is also a convenient way of moving collaborators in the virtual environment by manipulating their 3D representation in the WIM, as shown by Beck et al. [Bec+13] (Figure 2.4-a). To support group navigation, Weissker et al. [WBF20] designed a teleportation technique that allows two users to travel together while adjusting their spatial formation. Later, they extended this technique to support groups of up to ten users [WF21] (Figure 2.4-b).

2.4 SUMMARY

This related work section defined large interactive spaces by providing a selection of relevant work on wall-sized displays, immersive virtual reality systems and augmented reality spaces. It also described the main characteristics of the interactive systems that I used in my research.

Moreover, I presented examples that illustrate interaction and collaboration aspects in such large interactive spaces. Some previous work introduced interaction techniques for managing a large amount of 2D content on wall-sized displays, while others proposed techniques for navigating and interacting with 3D content in mixed reality environments. Although the content differs between these two contexts, there are similarities in terms of interaction, especially regarding pointing techniques or other techniques to interact at a distance, such as portals or tangible props. These similarities arise from the fact that users interact within a 3D space in both cases. Nevertheless, it is worth noting that no standardized interaction exists in these systems and that techniques usually need to be customized to suit specific application contexts. It is also interesting to observe that only very few techniques offer true collaborative interaction among co-located users.

Finally, I surveyed previous work on remote collaboration across large interactive spaces, including telepresence systems and collaborative virtual environments (CVEs). For telepresence, some systems focus on accurately conveying gaze direction, while others provide users with the ability to share and interact with digital content. However, none of these systems deal with very large interactive spaces, where multiple users can both move freely and interact with shared content from various physical locations. For CVEs, some previous work has explored how avatars can enhance mutual awareness among remote users. While video can be valuable in certain collaboration contexts, only a few systems integrate video into mixed reality environment. It is also worth noting that a large body of work on CVEs has focused on interaction in collaborative virtual environments, allowing users to collaboratively manipulate virtual objects and navigate in the virtual environment. Telepresence systems and CVEs offer different advantages, depending on the collaboration contexts and application domains. Nevertheless, almost no previous work has attempted to combine the two approaches by connecting users of heterogeneous devices, including both immersive and non-immersive systems.

3

FROM INTERACTION TO COLLABORATION IN A SHARED INTERACTIVE SPACE

Large interactive spaces are powerful tools for fostering collaboration on both digital and physical content, as they can accommodate multiple users within the same system. In these spaces, users can easily perceive each other's activities, share information and distribute tasks. Nevertheless, providing all users with appropriate interaction capabilities is a fundamental prerequisite for an effective collaboration.

Although large high-resolution displays and mixed reality technologies are becoming more mature and widespread, interaction in such systems remains a challenge. On the one hand, novel interaction paradigms are needed to enable users to fully exploit the specific features of these technologies. In particular, the paradigms must handle large visualization spaces and 3D interaction, while allowing users to move freely within the system and interact from multiple locations. On the other hand, these paradigms need to be designed considering the collaborative nature of the interaction from the outset. They must enable all users to act together with similar capabilities, while managing conflicts that may arise when users interact with the same content and occupy a shared space.

In this chapter, I present my research on interaction and co-located collaboration within a shared interactive space. This work addresses three fundamental aspects of large interactive spaces by investigating (i) how users can handle a large visualization space, (ii) how they can interact in a 3D space, and (iii) how they can take advantage of the large physical space surrounding them. This chapter is divided into three sections, each covering a different aspect. For each aspect, I first propose interaction paradigms that leverage the specific features of the system and support multiple users. I then study how these paradigms impact collaboration and how they can be extended to further enhance collaborative interaction. This chapter emphasizes the application of the proposed interaction paradigms to various collaborative design scenarios, including 3D sketching, computer-aided design (CAD) and industrial assembly tasks. Although it provides domain-specific solutions for managing complex data, facilitating 3D interaction and enhancing collaboration, these paradigms can be adapted to a wide range of contexts and data types.

3.1 HANDLING THE LARGE VISUALIZATION SPACE

Ultra-high-resolution wall-sized displays can present large amounts of visual information with a high level of detail, as presented in Section 2.1.1. However, interacting with such large visualization spaces requires specific techniques that provide a wide range of actions along with a high degree of precision. Touch interaction is a relevant option for interaction with wall-sized displays, as it satisfies these requirements while also enabling multiple users to interact simultaneously from different locations. Nevertheless, touch interaction techniques need to be redesigned to suit wall-sized displays, as standard touch-based techniques are mainly designed for single users with small handheld devices or horizontal screens.

While most previous work about interaction with wall-sized displays concentrates on 2D content (Section 2.2.1), the benefits of wall-sized displays can be extended to 3D data. This section targets a collaborative scenario in which a design team wants to create and explore numerous alternatives of 3D computer-aided design (CAD) objects on a wall-sized display. The first subsection focuses on the design of touch-based interaction to manipulate 3D objects on a vertical display. This work was published at VRST 2016 [**LF16**]. The second subsection investigates how touch interaction can be used to generate and distribute multiple design alternatives of a CAD object on a large display. Building on these outcomes, the last subsection presents a collaborative system that supports collaborative exploration of CAD data on a wall-sized display, and evaluates its benefits. This last study leads to more generic recommendations on how a large visualization space can be shared to enhance collaboration and empower users to perform complex tasks. The work described in these last two subsections was published at CHI 2020 [**Oku+20**].

3.1.1 *Multi-touch 3D interaction on a vertical display*

We explored the design of touch-based interaction for users interacting with 3D content while standing in front of a large wall-sized display. This scenario is motivated by the increasing availability of large multi-touch screens in meeting rooms, classrooms or public spaces. Multi-touch interaction provides a relevant solution for manipulating 3D objects in such contexts, as it can be easy to use and learn even for non-experts. It also does not require additional hardware beyond what is already embedded in the device.

Previous work has investigated touch-based 3D interaction on standard devices such as smartphones, tablets or tabletops. It proposes several solutions to perform 6-degree-of-freedom manipulation of 3D objects, such as mimicking direct 3D manipulation on the objects [RDH09; HCC09; JSK12] or combining different touch inputs to control separated degrees of freedom (DOF) [HCC07; MCG10a]. Other studies demonstrate that controlling separated DOF [HCC07] or separating the control of translation and rotation [MCG10b] improves the performance of the 3D manipulation. While these outcomes are still valid for wall-sized displays, the proposed techniques must be adapted to the new constraints introduced by such devices. In particular, touch input techniques that require the use of several fingers from the same hand may not be convenient when users need to perform actions at the top or bottom of a large wall-sized display. Additionally, long drags across the screen should be avoided to prevent user fatigue.

Other studies have explored touch interaction on large displays for 3D navigation [Yu+10] or 3D data exploration [Lop+16; Cof+12]. However, most of the proposed techniques consider that users stay static in front of the display or require control devices such as tablets or tabletops. In the context of meeting rooms, classrooms or public spaces, we want to design solutions that do not require these additional devices and instead rely solely on direct interaction with the display.

We designed In(SITE) [**LF16**], an Interface for Spatial Interaction in Tactile Environments, to explore touch-based 3D interaction on wall-sized displays. This technique combines bimanual touch interaction and object teleportation features to enable users to perform 6-DOF manipulation on a large vertical display. In(SITE)

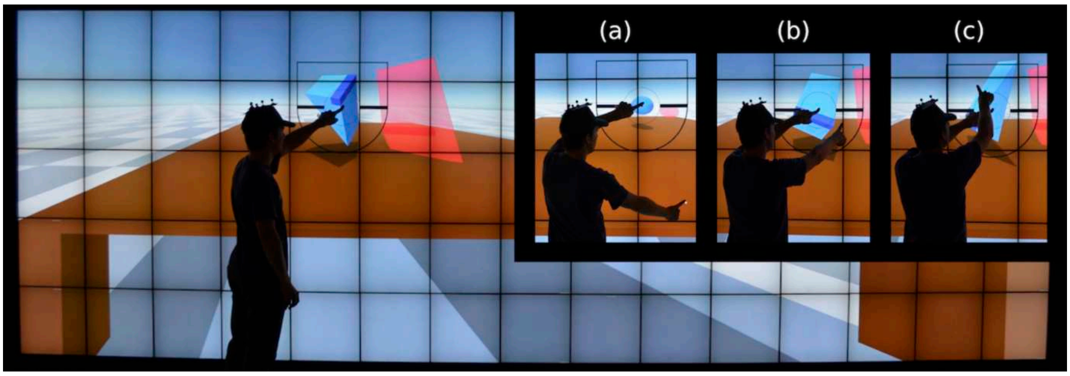


Figure 3.1: 6-degree-of-freedom manipulation of a 3D object on a multi-touch wall-sized display with the In(SITE) technique: the user performs (main picture) `x and y` translations, (a) `z` translation, (b) roll rotation, and (c) pitch and yaw rotations.

focuses on the selection, translation and rotation of 3D objects, but does not include scale modification in this first version.

In(SITE) provides a widget divided into several areas that enable separate manipulation of the different DOF (Figure 3.1). When users touch the screen, a raycast is performed in the 3D virtual environment starting from their head position and passing through the touch point on the screen. If the ray hits an object, the object is selected and the widget appears under the finger if it stays in contact with the screen for at least $1s$ (long touch). Users then control the x and y translations in a plane parallel to the screen by moving this primary finger. The z translation is controlled by touching outside the widget with any finger of the other hand. As this secondary finger moves closer to the primary finger, the object moves closer according to the ray axis, and vice versa. This interaction is inspired by the *Z-technique* [MCG10a]. For rotation, the lower area of the widget allows users to control the *roll* with the secondary finger by doing curved gestures, following the object rotation. The upper area allows users to manipulate the *yaw* and *pitch* with the secondary finger by doing respectively horizontal and vertical movements. The *yaw* and *pitch* rotations can be combined by performing diagonal movements.

To avoid long drags across the screen, In(SITE) includes teleportation features to achieve object translation over large distance. Users first need to select an object with a short touch (less than $1s$), instead of the long touch used to display the widget. The object color is changed to provide feedback that it has been selected. Once users have selected an object, they can teleport it anywhere in the virtual environment using two methods:

- A short touch at the destination, either on the floor or on another object, makes the object fall from above the destination. This method is particularly useful for virtual environments with physical simulation, as the object can be stacked on top of other objects.

- A long touch at the destination makes the object appear under the finger, at the location defined by the intersection between the ray and the virtual environment. The interface then switches to manipulation mode and displays the widget, allowing users to perform final position adjustments. This method is especially useful when users want to reach a precise position for the object.

We conducted two controlled experiments to assess the usability and performance of In(SITE) in comparison to a standard virtual ray technique, also known as the ray-casting technique [JFH94; Min95]. We selected this technique as a baseline because it is widely used in many virtual reality applications and particularly relevant for moving objects over long distances in large visualization spaces. To overcome certain limitations of the virtual ray technique for 3-DOF rotation and ensure a fair comparison, we augmented the technique with a feature that enables rotation along a vertical axis. Both experiments were performed on a $5.90m \times 1.96m$ wall-sized display (see description of the WILDER system in Section 2.1.1). This wall-sized display does not support stereoscopic vision, but we implemented motion parallax to improve depth perception. Both experiments involved 16 participants each, and focused on a docking task with targets positioned on the floor or in mid-air. We hypothesized that the virtual ray would be faster for translation, whereas In(SITE) would be more accurate for fine adjustments, especially when rotation is involved. We also predicted that the teleportation can improve both techniques for translation.

The first experiment used spheres for the docking task, assessing only translation. It compared In(SITE) and the virtual ray technique, both with and without teleportation. The results did not show a significant effect of the techniques on the task completion time and the mean values were almost similar, suggesting that participants reached close levels of performance with both techniques. However, In(SITE) led to significantly fewer overshoots than the virtual ray while adjusting the final position of the object. This is confirmed by the subjective questionnaire which reported that participants found In(SITE) easier to use and more precise. In addition, this questionnaire showed that participants preferred both techniques with teleportation and considered them easier to use and less tiring than the ones without teleportation.

The second experiment used edges, including rotation in the task. It compared In(SITE) with the virtual ray technique, but did not include teleportation as the task involved only short translation. The results did not reveal a significant effect of the techniques on the task completion time, but In(SITE) also led to significantly fewer overshoots than the virtual ray for this task.

Overall, these experiments suggest that In(SITE) can be an alternative for interacting with 3D content on wall-sized displays, as participants reached close levels of performance and better precision for fine adjustments with In(SITE) compared to a standard virtual ray technique. According to participants' feedback, the teleportation feature improves translation tasks in terms of ease of use, fatigue, and user preference. However, In(SITE) is a first prototype, which can be further improved. In particular, additional work would be required to investigate other designs of the widget, adjust transfer functions for indirect interaction and adapt the technique to stereoscopic display by following the guidelines presented by Valkov et al. [Val+11].

3.1.2 *Interaction with numerous design alternatives*

In the context of industrial design, we focused on using touch interaction to explore a large number of design alternatives on a wall-sized display. The first objective was to enable non-experts in computer-aided design (CAD) to modify such parametric models and to generate many design alternatives, using simple touch interaction instead of specifying complex geometric parameters. The second

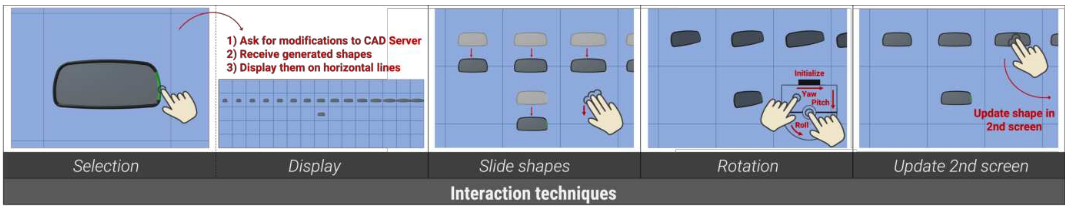


Figure 3.2: Interaction with *ShapeCompare*: when users select a part (Selection), the system displays a set of design alternatives on a row of the screen (Display). All alternatives can be scrolled up and down with a three-finger drag (Slide shapes) and rotated in 3D with the In(SITE) widget (Rotation). A specific alternative can be selected with a two-finger long press (Update 2nd screen) and displayed in context on another screen.

objective was to provide them with a solution for distributing and comparing these design alternatives by taking advantage of the large visualization space available. The target scenario was a co-located collaborative design situation, where a multidisciplinary design team, including designers, engineers, and ergonomists, wants to evaluate and adjust product designs using digital mock-ups, as described by Mujber et al. [MSH04].

Several interaction techniques have been proposed to assist designers with drawing and sketching during the early stages of the design process, including immersive drawing [Isr+09; SH16], surface modeling [Fio+02], digital tape-drawing [Bal+99; Gro+01; Fle+04; KZL07] and rapid prototyping with bimanual interaction [Ara+13]. However, only a few techniques target detailed design stages requiring the modification of parametric CAD models. A CAD model is a solid model defined by a set of mathematical operations (e.g., extrusion and boolean operations) applied to 2D sketches. Unlike drawing or surface modeling, modifying CAD models requires to manipulate parameters, which necessitates extensive training. Although some solutions enable non-experts to modify CAD data [Mar+17; Cof+13], they are limited to a single CAD model and do not support the generation of new alternatives.

We designed *ShapeCompare* [**Oku+20**] to meet the following criteria: (i) interaction in a large space, (ii) native CAD data modification and (iii) multiple-design comparison. We first implemented a service which generates multiple alternative shapes by varying parameter values of a native CAD model. This service can load native CAD files, modify parameters on request, and send back tessellated meshes using the CAA API of CATIA V5[1]. It also maintains a direct link between the 3D mesh parts and the CAD parameters using a *labeling* concept [CB04].

We created a first prototype to generate and visualize new design alternatives on a wall-sized display. For shape generation, users touch the part they want to change on a displayed shape (Figure 3.2). If the part can be modified, it turns green and the system requests a set of new alternatives by varying the CAD parameter related to this part. In this first version, we defined a minimum and maximum parameter value for each part, and chose a predefined number of values equally distributed within the range. For visualization, new shapes are distributed on an entire row of the screen, above other versions of the CAD model. Each row represents a set of design alternatives for a specific modification. Users can scroll up or down the alternatives using a three-finger interaction, and thus see the full

---

[1] `https://www.3ds.com/products-services/catia/`

Figure 3.3: User study setup: the target shape is displayed with a transparent yellow color in a realistic environment on an external screen next to the wall-sized display.

design history. Users can also select a part of any shape in the design history and restart modification from that shape. To handle 3D objects, users can rotate the alternatives by using the widget provided by In(SITE) (Section 3.1.1). The rotation of all alternatives is synchronized to maintain a similar same viewing angle.

To assess this first prototype, we conducted a user study and brainstorming session with five students from the civil engineering department of our university. Although they were not CAD experts, they had knowledge of parametric modeling and design process. They had to modify a car rear-view mirror with *ShapeCompare* to reach a given target shape within a 5-minute time limit. They could select a specific alternative with a two-finger long press on the wall-sized display and visualize it in an automotive cockpit on an external screen (Figure 3.3). The target shape was overlaid with a transparent yellow color on this external screen, simulating the design skills of experts assessing alternatives in a realistic environment.

We evaluated our design through the observations of participants' behaviors and interviews. Overall, participants appreciated the interaction techniques and found the system to be beneficial for novice users as it does not require understanding or manipulating parameter values. They also found the shape visualization nice and helpful in generating new ideas. All participants agreed that, although *ShapeCompare* has limited functionalities and cannot replace traditional CAD software, it is valuable for adjustments that do not require changing the entire design intent.

The study outcomes helped us identify the main issues that the participants faced. Firstly, all of them found it difficult and frustrating to understand how part selection affects shape deformation. Secondly, they often needed time to find out how the generated shapes on the new row are different from the one they selected. Because the parameter values used to generate shapes are always distributed between a fixed minimum and maximum, the initial shape on the new row is not displayed above the previously selected one, but at a random position. Aside from these issues, participants often had difficulty distinguishing differences between neighboring shapes, especially for the radii of corners, top, and bottom parts.

Based on these results, we then redesigned *ShapeCompare* to improve: (i) understanding of shape modification and (ii) visualization of design history. For the first aspect, we drew inspiration from the *Suggestive Interface* [IIH07] and created a widget that shows small thumbnails presenting the minimum and maximum shape modification for all parameters of each part (Figure 3.4-left). This widget becomes visible when users select a shape, and allows them to choose a thumbnail for generating the corresponding set of design alternatives. For the second aspect, we changed the way the system generates design alternatives to ensure that the selected shape always appears in the middle of the new row, with equal numbers

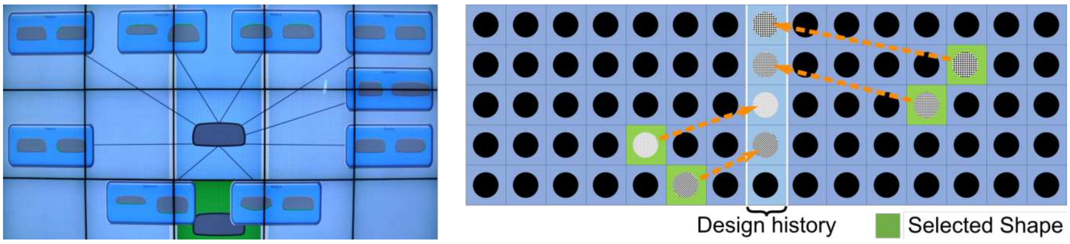


Figure 3.4: Updated version of *ShapeCompare*: (left) the selection widget shows small thumbnails with minimum and maximum shape modifications for all parameters and (right) all selected shapes are displayed in the middle of the next row to improve the visibility of the design history.

of alternatives displayed on both its left and right sides (Figure 3.4-right). Instead of defining a fixed minimum and maximum for the parameter values, we defined a specific offset for each parameter. The system thus generates the shapes by incrementally increasing and decreasing the parameter by the offset. Consequently, the middle column gathers all the previously selected shapes, allowing users to easily track the progression of modifications.

In summary, we employed an iterative design process involving potential users to create a custom interaction technique that facilitates CAD model exploration for non-experts. Our objective was to challenge the traditional design methodology in which users iterate on a single model. Instead, we proposed a solution that allows users to generate many design alternatives and explore them on a wall-sized display. Although we studied this approach in a specific context, visualizing "small multiples" on a wall-sized display could be extended to other contexts as long as parameter variations are involved. For instance, it can be applied to generative design [Che+18; Kaz+17] in which users can specify preferred designs to an Artificial Intelligence, or physical simulations such as weather predictions in which users can run several simulations with varying parameter settings. Furthermore, our approach is valuable for collaborative design, as it allows multidisciplinary teams, including non-CAD experts, to explore, compare, and reflect on design alternatives. I detail such a co-design scenario in the following subsection.

3.1.3 *Collaborative data exploration on a large display*

We aimed to investigate the potential benefits of using a wall-sized display to enhance collaboration within design teams during review meetings. Most of the time, the design process is iterative and relies mainly on two steps that involve many stakeholders: design discussion and CAD data adjustment. We aimed to create a collaborative system using a wall-sized display that could merge these two steps. It must enable multidisciplinary teams to collectively share and organize the large visualization space for generating and comparing numerous design alternatives.

Review meetings are a crucial aspect of the industrial design process. They typically take place in interactive systems that provide a full-scale design visualization, a large interactive space and a collaborative environment. For example, *Portfolio Wall* [Bux+00] displays different designs as tiled thumbnails on a large screen, mimicking traditional wall-mounted corkboards. Khan et al. [Kha+05] propose a tool that highlights the area where users need to pay attention on a projected display,

Figure 3.5: *ShapeCompare* enables multidisciplinary experts to generate and explore a large number of design alternatives on a wall-sized display.

thereby facilitating group meetings. Additionally, several virtual reality systems can display CAD data [Ber99; Ran+01; Rap+09]. However, all these systems limit users to comparing just a few static design alternatives during each review meeting, and these alternatives usually need to be prepared beforehand. Consequently, designers are unable to explore new ideas by generating and modifying design alternatives of CAD data in real time during the meeting. This limitation hinders their creativity and forces them to rely on a time-consuming iterative process. A few systems enable users to modify native CAD data [Mar+17; **Oku+18a**], but they focus on deforming a specific CAD model and do not consider the generation of new alternatives.

We used *ShapeCompare* [**Oku+20**] to create a collaborative system enabling multiple users to generate a large number of design alternatives, distribute them on a large wall-sized display and collaboratively compare them (Figure 3.5). This system relies on the interaction techniques described in Section 3.1.2. In such context, the wall-sized display is an efficient tool to show multiple variations of a same object, foster design discussions among multidisciplinary experts and enable them to explore more alternatives without using a conventional CAD system.

We conducted a controlled experiment comparing *ShapeCompare* with another visualization technique suitable for standard screens, called *ShapeSlide*. *ShapeSlide* displays only one shape at a time and enables users to change the shape displayed at the center of the screen with a sliding gesture. We used the same wall-sized display for both conditions (see description of the WILDER system in Section 2.1.1) in order to reduce bias that could be introduced by different devices, participants' positions, or interaction techniques. However, only a small part of the wall-sized display was used for *ShapeSlide*, simulating the use of a smaller screen. 12 pairs of participants performed a constraint solving task with both conditions. This task was based on actual industrial practices and involved modifying a car rear-view mirror, simulating expert negotiation on various design criteria. Due to difficulties in accessing actual industrial designers, we controlled participants' expertise by giving them individual design criteria based on simple numerical values computed by the system. In each pair, the first participant focused on the general properties of the mirror shape, such as aspect ratio and asymmetric balance, while the second concentrated on the mirror reflection, including visibility and size of the reflective area. The task ended when the design satisfied each participant's criteria, and when they both agreed on it. We hypothesized that participants would find the right design faster and with fewer iterations with *ShapeCompare* than with *ShapeSlide*.

The main results show that pairs of participants reached the right design significantly faster with *ShapeCompare* than with *ShapeSlide*. The questionnaires also highlight that *ShapeCompare* was perceived as more helpful for communicating with the partner, and generally preferred by participants. The smaller task completion time with *ShapeCompare* could be explained by this better communication between participants. This is supported by the significantly larger number of deictic instructions used by participants with *ShapeCompare* than with *ShapeSlide*. The alternatives of *ShapeCompare* were often used as references for communication and help participants convey their ideas. On the contrary, more words related to *Magnitude* were used with *ShapeSlide* (e.g. *"much more"* or *"a bit less"*), as they needed to describe their requirements verbally or with their hand gestures. It demonstrates that the alternatives of *ShapeCompare* help collaborators build a common ground, as defined by Clark [CB91], and thus minimize communication costs. Finally, the results from the NASA TLX questionnaire did not show significant differences between conditions, which suggests that displaying lots of alternatives with *ShapeCompare* does not substantially increase the cognitive load of participants.

In summary, this work investigates co-located collaboration on a wall-sized display in the context of industrial design. We used *ShapeCompare* to create a collaborative system that enables multiple users to generate numerous alternatives of a CAD model and distribute them on the wall-sized display. In a controlled experiment, we demonstrated that visualizing many alternatives on a wall-sized display enhances design exploration and negotiation by increasing the common ground among collaborators. These findings can be extended to more generic contexts involving the comparison of multiple alternatives. In particular, the concept of “small multiples” holds promise for facilitating multidisciplinary teams in collaboratively exploring, discussing, and reflecting on their ideas using a wall-sized display. The current system remains a research prototype, which leaves plenty of space for exploration and improvement in terms of visualization and interaction. For instance, investigating additional methods for classifying and merging relevant design alternatives is a potential direction for future research.

3.2 INTERACTING IN A 3D SPACE

Large interactive spaces enable users to interact in a 3D space, as most of them can detect user positions and gestures with advanced tracking systems, such as infrared or "Inside-Out" tracking systems. 3D interaction has long been employed in mixed reality for interacting with 3D content in virtual environments (Section 2.2.2). However, it can also be valuable even if the content is visualized on 2D displays. For example, 3D interaction with a physical prop representing a brain was used to control brain scans displayed in 2D on a wall-sized display [Bea11]. Moreover, 3D interaction offers many opportunities in collaborative contexts, supporting multiple users interacting in the same space and providing them with their own interaction area. It can also allow collaborative interaction as users can manipulate together virtual objects in the 3D space [ADL10b; ADL10a]. Despite its advantages, 3D interaction needs to be adapted to its application contexts, as no standards yet exist. Designers should also consider that it may be imprecise and tiring for users if no precautions are taken. When multiple users interact in the same 3D space, sharing the space may not be obvious, potentially leading to conflicts.

In this section, I explore 3D interaction in various design scenarios. The first subsection focuses on 3D interaction for modifying parametric CAD objects in the context of industrial design. The objective is to enable non-CAD experts to perform direct physical actions on the 3D shape of CAD objects in immersive virtual reality systems. This work was published in *Frontiers in Robotics and AI* [**Oku+18a**] and in a book chapter [**Oku+21**]. The second subsection investigates collaborative sketching in augmented reality. It presents a system that allows several users to interact with multiple versions of 3D content in the same physical space, managing conflicts and fostering creativity. This system was published at IHM 2023 [**FFT23**].

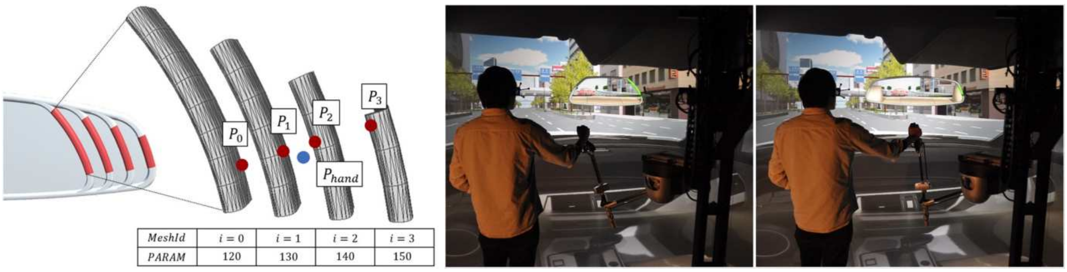
3.2.1 *CAD object deformation with physical actions*

We focused on techniques to modify parametric CAD data in large immersive VR systems. Our goal was to allow users to move around 3D CAD objects, feel them through haptic feedback, and modify them by performing physical actions on their surface. We targeted a scenario where non-CAD experts, such as stylists or designers, want to make simple modifications to CAD objects during a product review session in an immersive system.

While it is possible to create and modify primitives and meshes using shape-based interaction [Fio+02; De +13], applying these interaction techniques to CAD data is challenging due to the unpredictable object deformation resulting from parameter changes. A few VR-CAD applications enable users to modify native CAD data in an immersive system [Bou+10; Mar+17], but they do not support direct interaction with the CAD object shape. For instance, Martin et al. [Mar+17] use a one-dimensional horizontal gesture to increase or decrease a parameter value.

We developed *ShapeGuide* [**Oku+18a**; **Oku+21**], which enables users to deform CAD objects by directly pushing or pulling object surfaces in the virtual environment (Figure 3.6-right). It can include haptic feedback to enable users to feel the CAD object shapes and to increase the precision of deformation actions in the 3D space. To begin the modification process, users must first select the specific part of the CAD object that they want to modify. To handle the “unpredictability” of the shape deformation when modifying CAD parameters, a dedicated service computes a large number of possible shapes from a set of discrete parameter values associated with the selected part (Figure 3.6-left). This mesh pre-computation introduces a loading time after the selection to ensure real-time interaction later. In our prototype, this operation takes a few seconds, but this time can be significantly reduced with more powerful hardware and parallel mesh generation. Once the system has generated the set of shapes, users can explore them using a 3D hand motion. The system computes the distance between users’ hand position and the nearest point on each generated mesh. It thus displays the closest mesh to the hand. If provided, haptic feedback is computed as an attractive force to the nearest point of each generated mesh using a magnetic force inspired by [Yam+02]. This haptic feedback attracts the user’s hand to the surface of the closest mesh, keeping the hand steady on one mesh or guiding it toward neighboring meshes.

In a controlled experiment, we compared *ShapeGuide* to the one-dimensional horizontal scroll technique previously used by Martin et al. [Mar+17]. The direction



MeshId	$i = 0$	$i = 1$	$i = 2$	$i = 3$
PARAM	120	130	140	150

Figure 3.6: (Left) *ShapeGuide* precomputes several meshes of a rear-view mirror and selects the displayed shape according to the user's hand position P_{hand}. (Right) *ShapeGuide* allows users to modify the rear-view mirror shape using physical actions with haptic feedback, while immersed in a virtual car cockpit.

of the scroll was static and may not be consistent with the shape deformation in most cases. We also evaluated the effect of haptic feedback assistance on both techniques. 16 participants had to deform a car rear-view mirror to reach a target shape displayed in transparent yellow. The experiment was conducted in a CAVE system (see description of the EVE system in Section 2.1.2). Participants could interact everywhere in the CAVE with a *Virtuose* haptic arm mounted on a *Scale1* carrier. When no haptic feedback was provided, the haptic device let participants interact with zero force and resistance. We hypothesized that participants would perform the task faster and be more likely to start the deformation in the correct direction with *ShapeGuide* than with the scroll technique. We also predicted that the haptic feedback would improve the accuracy of both techniques.

Results demonstrate that *ShapeGuide* is 42% faster than the scroll technique for the rear-view mirror deformation. This improvement can be explained by a better consistency between shape deformation and user hand motion. In particular, we observed that *ShapeGuide* reduces by 80% the chance that participants move their hands in the wrong direction at the start of their gesture. Additionally, participants perceived *ShapeGuide* as less mentally demanding, less frustrating, less difficult to use, and preferred it over the scroll technique.

The main limitation of *ShapeGuide* is that it tends to produce more overshoots than the scroll technique, especially for parts where the shape variations are close to each other in 3D space. An overshoot occurs when participants reach the target shape but continue their gesture beyond it, causing them to move to the next shape and then come back. However, the results also show that haptic feedback reduces the number of overshoots for both techniques. Therefore, it can be an effective solution to improve the precision of *ShapeGuide*.

Overall, *ShapeGuide* provides an effective solution for deforming CAD objects within an immersive VR system, enabling physical actions to be performed directly on the object 3D shape. It can enhance the current industrial design process by allowing non-CAD experts to modify CAD objects without requiring an in-depth understanding of the CAD data internal organization. This will help avoid time-consuming iterations and potential misunderstandings that can occur when designers have to request modifications from CAD engineers. However, further work is still needed to improve mesh generation. It would be important to reduce generation time and allow users to select the number of generated meshes and the scale level of CAD parameter changes. Additionally, further evaluation of *ShapeGuide* with other industrial CAD models would be necessary.

3.2.2 *Collaborative sketching in augmented reality*

Large augmented reality spaces are valuable solutions for collaborative design, enabling co-located users to create virtual content that overlays their shared physical space. However, conflicts arise when several collaborators want to add or modify virtual content around the same physical objects. Although sharing content among collaborators is crucial in the creative process [Wal+20], others' content can sometimes distract users and hinder their creativity [GBR12]. Our objective was to create a system that helps multiple users share an augmented reality space, allowing them to independently develop their own virtual content while remaining aware of each other's activities and productions.

A wide range of research work focuses on co-located collaboration in mixed reality. Typically, users see and interact with identical virtual content, but a few systems introduce the ability to access or switch between simultaneous versions of this content. In *Slice of Light* [Wan+20a], multiple learners are immersed in distinct virtual environments while being co-located in the same physical space. The system allows the teacher to switch between all the environments by moving in the physical space. *Photoportals* [Kun+14] propose creating portals to access different locations in time or space within the virtual environment. *Spacetime* [Xia+18] uses containers to store and manipulate multiple versions of virtual objects, avoiding conflicts during concurrent manipulation in virtual reality. *VRGit* [Zha+23] provides a tool similar to a version control system to manage various versions of a virtual environment, thereby facilitating collaborative editing. However, these systems only address virtual reality, and do not take into account the relationship between virtual content and physical space, which is a crucial aspect of augmented reality.

Among previous work related to augmented reality, Looser et al. [LBC04] introduce magic lenses that display different layers of virtual content. However, they focus on the technical aspects and do not explore how these layers could be useful for collaboration in a creative process. The concept of *Duplicated Reality* [Yu+22] proposes to duplicate a portion of the physical world into an interactive virtual copy located elsewhere in the augmented reality space. By annotating this virtual copy, a user can guide another user who is performing actions in the physical world without disturbing them. While this system prevents conflicts during interaction, it does not handle multiple versions of the virtual content.

We developed a conceptual framework [**FFT23**] for co-located collaboration in augmented reality. This framework targets design scenarios where collaborators use physical objects as context, landmarks or guides to create 3D virtual content. It allows multiple versions of the virtual content to be associated with a single physical object, potentially representing multiple design alternatives. Users have the freedom to independently control which versions they perceive, and create their own versions without being constrained by those of others. They can also decide to share or not their versions, depending on the stage of the design process.

We reify each version as a *Version Object*, following the concept of reification proposed by Beaudouin-Lafon and Mackay [BM00]. *Version objects* are interactive representations that take the form of semi-transparent spheres containing a preview of the related virtual augmentations (Figure 3.7-c). Users have the ability to grab *Version Objects* and move them into space. *Version Objects* can thus be grouped

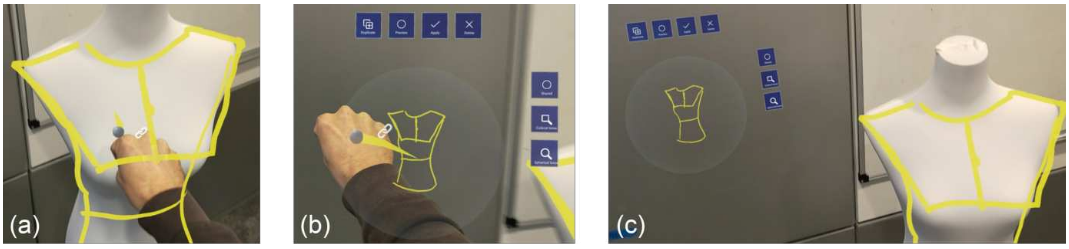


Figure 3.7: Creation of a *Version Object*: (a) a user grabs the 3D augmentations of a physical object. (b) This action creates a new *Version Object* with a preview of the augmentations. The *Version Object* can be moved and (c) stored in the AR space.

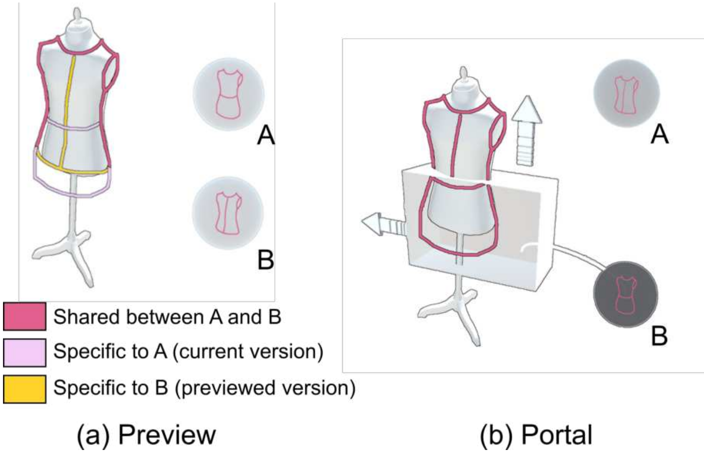
together in space, compared with each other, shared between collaborators and applied to the appropriate physical objects.

Users can create a *Version Object* by performing a grabbing gesture on a specific physical object at any time (Figure 3.7-a,b). This new *Version Object* represents the state of the object virtual augmentations at the time of its creation, similar to a photograph taken by a camera. By creating multiple *Version Objects*, users have the ability to capture various stages of their design process or save a version before making edits. When a *Version Object* is created, only its creator can see and access it. However, the creator can choose to share it with others.

Users can switch among different versions associated with a physical object by grabbing a *Version Object* and dropping it onto the corresponding physical object. This allows them to easily return to a previous version, explore various design options that they have created, or review versions shared by others. Our framework also provides users with the ability to simultaneously view multiple versions for comparing design alternatives. They can use either a preview that superposes two versions using transparency and color coding (Figure 3.8-a), or a 3D portal that renders one version inside the portal and another one outside (Figure 3.8-b). This portal can be freely moved in space by users.

To support collaborative design with *Version Objects*, the framework allows users to synchronize or desynchronize the virtual augmentations they see on a physical object. Desynchronization occurs when a user applies a specific *Version Object* to a physical object, thereby switching to a different version of virtual content compared to their collaborators. This feature can be useful to explore different design ideas or to benefit from a private space. During desynchronization, modifications made by collaborators are visible for a very brief period before fading out. This serves as feedback of collaborators' actions and indicates that the augmented reality space is temporarily out of sync. Users can then re-synchronize the virtual augmentations they see to work on the same content or share design ideas. Synchronization occurs when a user requests to synchronize with a specific collaborator. Modifications made by each user become visible to all. A preview mode also allows users to glance at a collaborator's version before deciding to switch to it.

We introduced a use-case scenario to illustrate the functionalities of our framework on a concrete example. This scenario involves two fashion designers who aim to create a new female jacket. They use augmented reality to sketch the virtual outlines of the jacket in 3D on a physical sewing mannequin, which serves as a support and guide for their creation (Figure 3.9). The two designers are co-located in the same room and use AR headsets. Our framework gives them the opportunity



(a) Preview (b) Portal

Figure 3.8: Comparison of two *Version Objects*: (a) a preview highlights differences between the current and previewed versions, or (b) a 3D portal allows users to explore differences between the versions inside and outside the portal.

Figure 3.9: Co-located users sketching in AR the virtual outlines of a jacket on a physical sewing mannequin.

to explore their own design ideas, to share them with each other and to collaboratively review them. The scenario consists of two phases: an initial divergence phase followed by a convergence phase, similar to what can be encountered in various creative or engineering processes. During the divergence phase, the designers use the desynchronized mode to individually create different design alternatives, but they can still share *Version Objects* to draw inspiration from each other. During the convergence phase, they switch to the synchronized mode to collaboratively review existing alternatives and make use of preview and portal tools to compare them. We implemented this scenario with two *Hololens 2* headsets from Microsoft.

In conclusion, we propose a framework that supports the various phases of collaborative design in augmented reality. It enables co-located users to perceive distinct versions of the virtual content associated with a physical object. These versions are reified into *Version Objects*, allowing users to control the virtual augmentations they see, explore their own design ideas, or share multiple design alternatives with collaborators. We illustrated the framework capabilities through a fashion design scenario, but its application can be extended to many design processes using augmented reality. Future work should focus on a formal evaluation of our framework, including its impact on the design process. In particular, the experience of co-located users viewing different content considerably differs from regular practices and can be disturbing. Therefore, further study is needed to understand how it influences collaboration and whether it increases users' cognitive load. Moreover, future work should consider more advanced solutions to display and organize the *Version Objects* in space. Drawing inspiration from the representations used by *VRGit* [Zha+23] in a virtual reality context could be valuable. Finally, we could consider creating virtual versions of physical objects using 3D reconstruction techniques, thus extending the concept of *Duplicated Reality* [Yu+22].

3.3 LEVERAGING THE LARGE PHYSICAL SPACE

By definition, large interactive spaces provide users with a vast physical space in which to move and interact. This physical space corresponds to the room or area accessible to users in mixed reality systems, but also to the space available in front of wall-sized displays or other 2D visualization systems. It offers valuable opportunities to explore virtual content through physical navigation. Previous work showed that physical navigation can improve spatial memory [JSH19] and performance in visual search tasks [BNB07] on 2D wall-sized displays. Additionally, physical navigation in mixed reality systems can provide users with vestibular cues that enhance spatial understanding [LaV+17] and immersion [Uso+99], while reducing cybersickness. When designing interaction in such systems, it is crucial to consider the physical space surrounding users and maximize physical displacements.

In collaborative contexts, a spatial relationship naturally exists among users who are co-located in the same physical space. This relationship needs to be considered during collaboration interaction or preserved in virtual environments. For example, when users equipped with VR headsets share the same room, preserving a consistent mapping between their physical and virtual positions allows them to have direct physical contact [Min+20] or co-manipulate shared physical props [SJF09] while immersed in the virtual environment.

This section mainly focuses on immersive virtual reality systems. The first subsection investigates several techniques that enable users to be aware of their physical space when navigating in a virtual environment. These techniques aim to optimize the mapping between the physical and virtual spaces, thereby maximizing users' physical displacements and enabling tangible interaction. These different results were published at EuroVR 2019 [**Zha+19**], VRST 2020 [**Zha+20**] and at the *Workshop on Everyday Virtual Reality* at IEEE VR 2021 [**Zha+21**]. The second subsection addresses a collaborative scenario in which multiple users independently navigate in a virtual environment while remaining in the same physical space. It proposes two collaborative navigation techniques that help users recover a consistent spatial mapping between their physical and virtual positions when they need to interact together. These techniques were published at IEEE VR 2022 [**Zha+22**].

3.3.1 *Virtual navigation with physical space awareness*

In the context of virtual reality systems, we aim to enhance immersion and interaction by taking advantage of the large physical space surrounding users. In all VR applications, this physical space, referred to as the users' physical workspace in this section, is mapped onto a specific region of the virtual environment. This spatial mapping allows users to physically walk in the virtual environment and to perform tangible interaction. Tangible interaction involves associating physical objects with virtual counterparts and using them as substitutes to manipulate the virtual content with passive haptic feedback [SVG15]. Such tangible interaction increases the sense of presence in the virtual environment [Hof98].

The mapping between the real and virtual world is a fundamental issue in every VR applications, and previous work has explored solutions to manage this relationship. Some applications [Che+19; Sra+16] opt to have a fixed one-to-one mapping between the real and virtual environments to avoid user collisions with

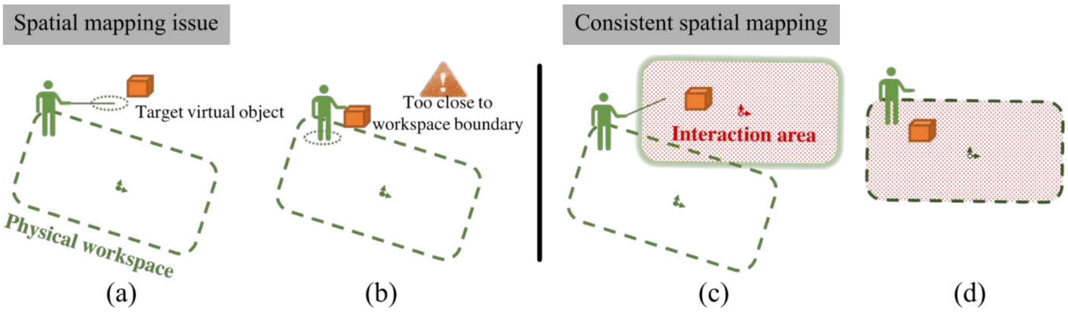


Figure 3.10: (Left) suboptimal spatial mapping that can occur after teleportation: (a) a user teleports themselves to interact with a virtual object without any knowledge of their position in the physical workspace and (b) the object is still out of the workspace boundaries after teleportation and cannot be reached. (Right) the application establishes a consistent mapping between the physical workspace and interaction areas: (c) a user teleports themselves to interact with a virtual object included in an interaction area and (d) the mapping is established during teleportation, allowing the user to walk and reach objects in the entire area.

the real world and grant direct access to all virtual objects. However, this approach constrains the size and shape of the virtual environment. To address this limitation, redirected walking [RKW01] or other view distortion techniques [SWK16] can be used to map a large virtual environment to a smaller physical workspace while allowing users to walk freely. Another approach is to use a self-overlapping architectural layout which allows users to walk through multiple virtual rooms while staying in the same physical room, as proposed by *Impossible Space* [Sum+12]. Nonetheless, these solutions may not be suitable for all applications since they require a reasonably large physical space, and physically walking could also be tiresome when users have to travel long distances.

Virtual navigation, such as teleportation, allows users to travel in a virtual environment beyond their physical workspace boundaries. Some previous studies propose taking into account the users' physical workspace as they navigate in the virtual environment, modeling it as a virtual vehicle [BT02] or a virtual cabin [**Fle+10a**]. However, virtual navigation alters the spatial mapping between the physical workspace and the virtual environment. This mismatch can result in a suboptimal mapping, causing users to unexpectedly reach the physical workspace limits, restricting physical walking and reducing direct access to virtual objects (Figure 3.10-left). Consequently, users may need to repeatedly rely on virtual navigation over short distances instead of using physical movements, which reduces immersion. Additionally, virtual navigation can break the relationship between tangible objects and their virtual counterparts. *Redirected Teleportation* [Liu+18] proposes to combine teleportation and physical walking by maximizing the space available for walking after each teleportation. To activate teleportation, users step into a portal to reposition and reorient themselves away from the physical space limits. However, this technique does not take into account the virtual objects that users need to access and cannot handle tangible objects.

In this section, we aim to recreate a consistent mapping between physical and virtual spaces in specific areas of the virtual environment after a virtual navigation. The main idea is that users can travel long distances freely by using virtual navigation without any distortion or need for additional actions, such as entering a portal.

However, when they need to interact with multiple virtual objects in the same area, the application can assist them in establishing a consistent spatial mapping (Figure 3.10-right). This enables users to directly access the objects by physically walking in the consistent area. Furthermore, the application can help users recover the spatial relationship between tangible objects and their virtual counterparts to perform tangible interaction. We focused on teleportation since it is widely used in VR applications, but this work can be extended to other navigation techniques. We considered standard teleportation with instantaneous transition and no viewpoint animation. Although it can lead to disorientation, this method is the most widely used in VR applications to avoid motion sickness.

We considered three solutions for defining the spatial mapping. First, application designers can specify the mapping in advance for specific areas when tasks are predefined, such as for virtual escape games or VR training with assembly tasks. For more generic applications, users can either choose the mapping manually or select automatically generated areas based on the layout of virtual objects. Finally, the mapping can be defined by the physical object positions for tangible interaction.

3.3.1.1 *Mapping defined by application designers*

As a first step, we investigated VR applications that involve predefined virtual object manipulations taking place in specific areas of the virtual environment. Application designers can thus position the interaction areas where manipulations will occur in advance. These interaction areas have the same dimensions as the users' physical workspace and will be used to create a one-to-one mapping between these physical and virtual spaces.

We introduced two switch techniques [**Zha+19**] based on teleportation to help users recover the mapping between their physical workspace and the interaction areas. In both techniques, users teleport themselves in the virtual environment by using a virtual ray to point towards the destination. However, when users point towards an interaction area, a specific representation is displayed to notify them that a special teleportation technique will be triggered. This teleportation technique adjusts the users' position and the orientation at the destination, ensuring their physical workspace matches the interaction area. Once users are teleported in this area, they can physically walk to access all virtual objects of the area.

The two switch techniques use different representations to display the interaction areas (Figure 3.11). The *Simple switch* shows a transparent cube with a green border indicating the boundaries of the area. The *Improved switch* uses the same cube representation, but it adds a semitransparent cylinder with a 3D arrow showing the users' future position and orientation in the area. This simplified avatar aims to help users anticipate their future location in the area and avoid disorientation. The avatar position and orientation are updated in real time, which means that users can see their avatar moving in the interaction area if they physically walk in their physical workspace before the teleportation.

We conducted a controlled experiment with 18 participants to compare the two switch techniques with a standard teleportation technique used as a baseline. The experiment was carried out using a CAVE system (see description of the EVE system in Section 2.1.2). Participants completed a box-opening task in four separate rooms connected by corridors. They traveled long distances between rooms using the teleportation technique. In each room, participants followed instructions to open

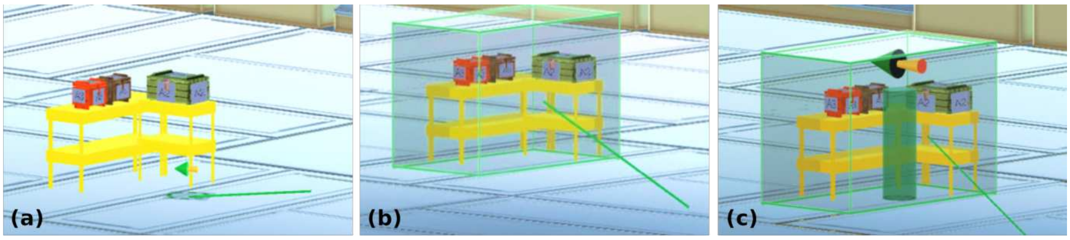


Figure 3.11: (a) standard teleportation technique: users' future position is displayed on the ground when they point towards the destination with a virtual ray. (b) *Simple switch*: a predefined interaction area is highlighted when users point towards it. (c) *Improved switch*: the representation included a semitransparent cylinder with a 3D arrow showing users' future position and orientation in the area.

three out of four boxes. The boxes were positioned in a U-shape in the room and were included in the same interaction area. Participants could teleport themselves either anywhere in the room with the baseline or within the interaction area with the switch techniques. We hypothesized that both switch techniques would improve task performance compared with the baseline. We also predicted that *Simple switch* would be faster, but would increase disorientation compared with *Improved switch*.

Results highlight that helping users recover a consistent spatial mapping improves performance and immersion. The two switch techniques significantly reduce task completion time, the number of teleportations required to achieve the task, and the collisions with physical workspace boundaries compared to the baseline. Participants also reported that both switch techniques were less mentally and physically demanding than the baseline. When comparing the two switch techniques, the *Simple switch* is faster than the *Improved switch* to perform the teleportation in the interaction area because users do not need to look at the avatar. However, the *Improved switch* seems to improve spatial understanding after teleportation as it reduces the time and head rotation required to find the first box, although we did not measure significant differences in the experiment.

In this work, we demonstrate the benefits of creating a consistent spatial mapping between users' physical workspace and specific regions of the virtual environment. This approach is particularly useful in complex scenarios that involve large-scale navigation and manipulation sub-tasks which require access to multiple objects in the same area. We also evaluated the effect of showing a simplified avatar to represent users' future location after teleportation in the interaction area. It seems that the avatar can be beneficial to reduce disorientation, even if it increases the time needed to trigger the teleportation by a few seconds. However, additional studies are required to fully assess the impact of the switch techniques on disorientation.

3.3.1.2 *Mapping defined by users*

Defining in advance the interaction areas where object manipulations will occur is not possible for all VR applications. In this second step, we explored more generic solutions that allow users to define by themselves the spatial mapping between their physical workspace and a designated area of the virtual environment. The goal is to make users aware that a virtual workspace related to their physical workspace exists and to enable them to choose the position of their future virtual workspace before each teleportation.

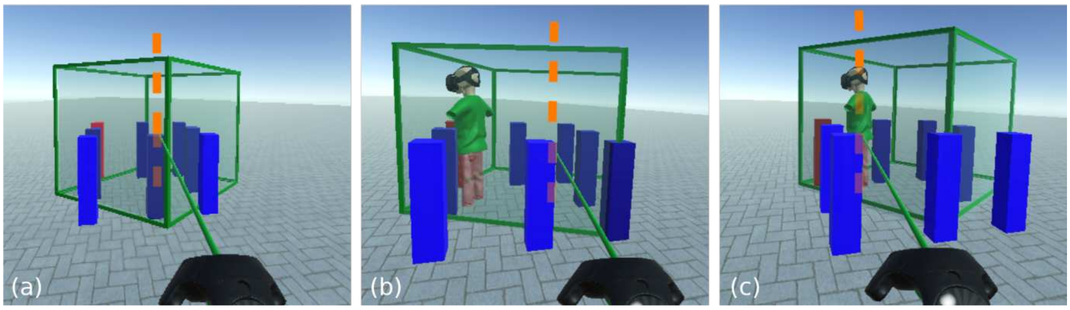


Figure 3.12: Three manual techniques allow users to position their future virtual workspace before the teleportation: (a) *Exo-without-avatar*, (b) *Exo-with-avatar* and (c) *Ego-with-avatar*. The orange dotted line represents the rotation axis of the 3D volume in this figure, but it is not visible to users in the virtual environment.

We designed both manual and automatic techniques to position this future virtual workspace [**Zha+20**]. In manual techniques, users directly adjust the position and orientation of a 3D volume representing their virtual workspace in the virtual environment by using a virtual ray attached to a VR controller (Figure 3.12). The intersection between the virtual ray and the virtual ground defines the 3D volume position, while a circular gesture on the controller touch pad controls its orientation. In automatic techniques, an algorithm computes a set of potential virtual workspaces according to the layout of the interactable objects in the virtual environment. Each virtual workspace alternative is represented by a 3D volume which becomes visible when users point towards it. Users can thus browse these alternatives and select the one they prefer with their virtual ray. In both types of techniques, the virtual objects inside the 3D volume are highlighted to help users understand what objects will be contained in their future virtual workspace. Once users trigger the teleportation, they are moved to the selected virtual workspace and can access all the virtual objects inside this workspace by physically walking.

In a first experiment, we compared three manual techniques (Figure 3.12):

- *Exo-without-avatar* implements an exocentric manipulation by allowing users to move and rotate the 3D volume around its central axis.
- *Exo-with-avatar* uses the same exocentric manipulation, but also includes a transparent avatar that shows users' future position in the virtual workspace after teleportation.
- *Ego-with-avatar* proposes an egocentric manipulation that also includes an avatar. It uses the users' future position (i.e., the avatar position) as the axis to move and rotate the 3D volume. This technique can thus be perceived by users in a different way: they move and rotate the avatar in the virtual environment, and the 3D volume just indicates the space that would be available after teleportation.

The purpose of this experiment was to evaluate the benefits of the avatar representation and to compare the exocentric and egocentric manipulations in terms of spatial awareness and performance. 12 participants performed a simple task in which they had to adjust their future virtual workspace position to enclose eight pillars, teleport themselves inside this new virtual workspace, and touch a specific pillar displayed in red. This last action was included to assess their spatial awareness. Participants used an *HTC Vive Pro Eye* VR headset in a $3m \times 3m$ physical area. We hypothesized that both techniques with avatar would reduce the time required to touch the pillar compared with *Exo-without-avatar*. We also predicted that *Exo-with-avatar* would reduce the time required to position the virtual workspace, but would increase the time required to touch the pillar compared with *Ego-with-avatar*.

Results show that both conditions with the avatar decrease the time required to touch the pillar after teleportation by over 50%, compared to *Exo-without-avatar*. On the contrary, the time spent positioning the virtual workspace before teleportation appears to be shorter without the avatar, which is consistent with our previous study (Section 3.3.1.1). Although the avatar slightly increases the time and cognitive load required to position the virtual workspace, it can help users better understand the upcoming teleportation and reduce disorientation. This finding is supported by participants' qualitative feedback, which reported better anticipation of their location after teleportation and less disorientation. Regarding the manipulation technique, the results did not show significant differences between exocentric and egocentric techniques in terms of user performance. However, participants preferred the *Ego-with-avatar* condition over the *Exo-with-avatar* condition since it was perceived as "*easier for positioning themselves*" and "*easier for finding*" the target pillar after teleportation.

In a second experiment, we compared a manual technique, an automatic technique and a standard teleportation technique in a more realistic task. Based on the results and participants' preference from the first experiment, we chose the *Ego-with-avatar* technique for the manual technique. For the automatic technique, participants used a virtual ray to select the future virtual workspace from a set of alternatives computed by the system, as described previously. 12 participants performed a task similar to an escape room, where they had to travel through 8 virtual rooms. In each room, they needed to grab 10 objects one by one and bring them in one of 2 boxes available in the room. Half of the rooms had a *No-overlap* layout, which consisted of 2 disjointed areas, each containing five targets and one box. The automatic technique computed 2 virtual workspace positions for this layout. The other half of the rooms had an *Overlap* layout for which the 10 objects and the 2 boxes were placed randomly in a single area. The automatic technique computed 4 overlapping virtual workspace positions to cover all this area. This experiment used the same VR setup as the first one. We hypothesized that both manual and automatic techniques would improve performance and sense of presence compared with the baseline. We also predicted that the automatic technique would perform better for *No-overlap* layouts and worst for *Overlap* layouts compared with the manual technique.

Results show that both manual and automatic techniques outperform the standard teleportation technique in terms of efficiency and immersion. In particular, they significantly reduce the task completion time, the number of teleportations required to achieve the task, and the collisions with physical workspace boundaries. Participants also reported a higher sense of presence in the IPQ questionnaire [RS02; Sch03] with both manual and automatic techniques compared to the standard one. Regarding the comparison between the manual and automatic techniques, both achieve close performance, but each one has advantages depending on the virtual object layout. The automatic technique causes fewer collisions with the physical workspace boundaries in sparse environments (i.e., *No-overlap* layout), but induces a higher cognitive load for crowded environments (i.e., *Overlap* layout) compared to the manual technique.

Overall, this work highlights the benefits of allowing users to choose the position of their virtual workspace before teleportation. Depending on the virtual object layout, both manual and automatic techniques can be valuable. For manual techniques, exocentric and egocentric approaches perform similarly, but users tend to prefer the egocentric approach. In addition, including an avatar to show the user's future position can decrease disorientation and minimize the time required to locate targeted objects after teleportation. Further studies would be mandatory to assess the proposed techniques in other scenarios including different shapes and sizes of physical workspaces, various virtual object densities and other types of tasks. Automatic techniques could be enhanced by adjusting the clustering algorithm based on the specificity of these scenarios. Finally, the visual representation of virtual workspace can be improved to prevent overloading the user's field of view.

3.3.1.3 *Mapping defined by tangible object positions*

Finally, we studied how to recover the spatial relationship between tangible objects and their virtual counterparts after a virtual navigation. Tangible interaction is a simple and inexpensive solution to provide haptic feedback by associating virtual objects with real objects that share similar physical properties. For instance, a real chair can allow users to sit in the virtual environment or holding a closed umbrella can simulate the sensation of holding a virtual sword [SVG15]. This passive haptic feedback can improve the sense of presence in virtual environments [Hof98]. However, when users perform virtual navigation to travel beyond what is possible according to their physical workspace boundaries, the spatial relationship between tangible objects and their virtual counterparts is disrupted, and tangible interaction is no longer possible.

We explored three advanced teleportation techniques to recover the spatial relationship with a specific tangible object [**Zha+21**] (Figure 3.13). We proposed to teleport (i) the user, (ii) the virtual object, or (iii) both to a new position, while recovering their relative positions. To demonstrate this, we developed a first prototype involving a tracked physical chair that can be used to sit in the virtual environment. The chair has virtual counterparts which the user can interact with in the virtual environment. This prototype used an *HTC Vive* VR headset to immerse the user in the virtual environment and a *Vive Tracker* to track the chair position. For all techniques, the interaction steps are the same: the user first selects the virtual object involved in the tangible interaction (i.e., one of the virtual chairs in our example), then the user adjusts the teleportation destination if necessary and, finally, the user triggers the teleportation.

In *user teleportation*, the selected virtual object serves as an anchor for the teleportation: the user's future position will be defined by the position relative to the object physical counterpart (Figure 3.13-a). For complex objects, such as the chair, the user has only one option as future destination. However, for more symmetric objects, the user may be able to choose several destinations around the object. For instance, in the extreme case of a ball, the user can have an infinite number of destinations all around the ball. According to the available options, sometimes the user does not have other choices than teleporting themselves inside other virtual objects, such as the table next to the chair, for example. In such cases, the system displays the user's

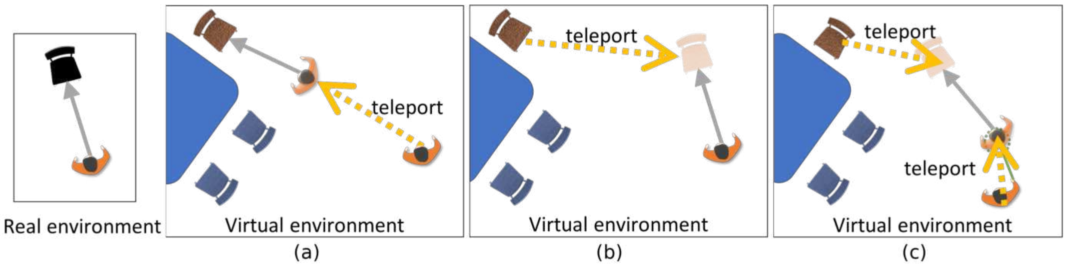


Figure 3.13: Three teleportation techniques allow tangible interaction by recovering the spatial mapping between a real chair and its virtual counterpart: they involve teleporting (a) the user, (b) the object or (c) both to a new position.

future position with a colored avatar, highlights the collisions with virtual objects and asks the user to physically move to a new position that avoids collisions before triggering the teleportation. Colored feedback shows both virtual objects colliding with the user's future position and available areas where the user should move to avoid these collisions.

In *object teleportation*, the user's current position serves as an anchor for teleporting the selected virtual object (Figure 3.13-b). This virtual object will be moved close to the user at the same relative position as the corresponding physical counterpart, enabling tangible interaction. As a consequence, there is a unique position possible for the selected virtual object. However, this position may already be occupied by other virtual objects surrounding the user. In such cases, the system detects the collisions in advance and computes a new position to which the user must teleport themselves before completing the object-based teleportation operation.

In *hybrid teleportation*, the user specifies where both themselves and the selected virtual object will be teleported (Figure 3.13-c). The user can use a virtual ray to define this future position, as in standard teleportation techniques. The virtual object position is computed based on the relative position of its corresponding physical counterpart. A specific representation at the intersection between the virtual ray and the virtual floor shows the future positions of both the user and the virtual object. The user can rotate this representation to adjust the future virtual object location all around their future position. This allows the user to avoid potential collisions with other virtual objects and to choose an appropriate position to prepare for the upcoming tangible interaction.

These three approaches have their respective advantages and disadvantages, and can be applied in different scenarios based on the tangible object characteristics. On the one hand, *user teleportation* may be more appropriate to allow users to access virtual objects that are considered immovable. On the other hand, *object teleportation* may be more suitable for interacting with small objects or tools at specific locations of the virtual environment. Finally, *hybrid teleportation* does not require additional strategies to prevent users or tangible objects from being teleported inside or behind other virtual objects. However, it can be time-consuming and mentally demanding for users. In future work, we need to conduct user studies to evaluate these three techniques in various contexts, including different shapes, sizes, and potential mobility of the tangible objects. Safety issues must also be considered in scenarios where tangible interaction alternates with free navigation. Indeed, physical objects can become invisible obstacles in the users' physical workspace when they no longer have a consistent spatial mapping with their virtual counterpart.

3.3.2 *Collaborative navigation to restore spatial consistency*

Multiple users can be co-located in a virtual reality system, for example, when they are all in the same room wearing VR headsets. In such cases, their relative positions in the physical space usually match those in the virtual environment. This spatial consistency enables users to have direct physical contact with each other [Min+20] or co-manipulate shared tangible props [SJF09]. However, this often excludes individual virtual navigation capabilities, such as teleportation, to preserve the one-to-one mapping between users' relative positions in the physical and virtual spaces. As a consequence, users can only explore virtual environments with approximately the same size and shape as their physical space. Some previous techniques allow co-located users to achieve virtual navigation, but restrict them to traveling together in the virtual environment. For example, *C1x6* [Kul+11] investigated group navigation in a projection-based system and *Multi-ray jumping* [WKF19] introduced collaborative teleportation techniques for co-located users wearing VR headsets [WKF19]. However, group navigation limits users' freedom during a continuous VR experience and is not suitable for many collaborative scenarios. To address these limitations, our objective was to design a system that enhances users' navigation freedom while preserving the capability of sharing the same physical space. In particular, this solution should enable users to independently navigate in a virtual environment, but also help them recover a consistent spatial mapping between their physical and virtual positions when they need to interact together. By doing so, this system would effectively support various phases of the collaboration, including individual exploration and tightly coupled manipulation.

Very few systems have explored how to restore the spatial mapping among co-located users after individual virtual navigation. The system proposed by Min et al. [Min+20] allows co-located users to individually explore a virtual environment larger than their physical workspace using redirected walking. When users need to perform direct physical interaction, such as shaking hands, they use a recovery algorithm that adjusts redirected walking parameters and recovers a consistent spatial mapping. However, this solution requires a large physical space and is not compatible with other navigation techniques, such as teleportation. In a single-user context, we proposed several techniques to recover spatial consistency between the user's physical space and a specific area of the virtual environment, as detailed in the previous subsection. The technique presented in Section 3.3.1.2 enables users to define the future position of their physical workspace in the virtual environment before a teleportation. We extended this technique to a collaborative context.

We proposed two techniques [**Zha+22**] that assist co-located users in recovering spatial consistency after individual teleportation when necessary for the subsequent collaborative interaction. Both techniques use a virtual representation of the users' shared physical workspace, which enables them to adjust the mapping between their physical and virtual spaces. We refer to this virtual representation as the "virtual workspace", as it corresponds to the area of the virtual environment that will be physically accessible to both users after teleportation. In addition, the future group configuration in the virtual workspace is represented by preview avatars, showing where users will be positioned after teleportation with a transparent color.

In the *Leader-and-Follower* technique, only one user, referred to as the *leader*, defines the future position of the virtual workspace before teleportation (Figure 3.14-left).

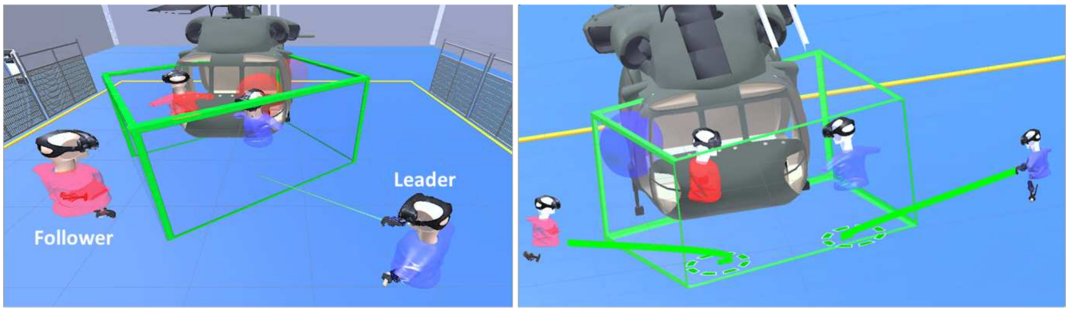


Figure 3.14: Two techniques for recovering a consistent spatial mapping after individual teleportation. The 3D volume framed in green represents the future position of users' virtual workspace mapped to their physical space. The preview avatars show their future positions inside this virtual workspace. (Left) the *Leader-and-Follower* technique allows one user to manipulate the position of the virtual workspace, while the second one can only communicate verbally regarding position requirements. (Right) the *Co-manipulation* technique integrates the inputs from both users, allowing collaborative positioning.

This user manipulates the virtual workspace with a virtual ray attached to a VR controller by using the *Ego-with-avatar* technique. We selected this technique based on the findings of a previous experiment described in Section 3.3.1.2. The virtual workspace position is defined by the intersection of the virtual ray with the virtual ground, while its orientation is controlled by performing a circular gesture on the controller touch pad. The rotation axis is determined by the future position of the *leader* within the virtual workspace. The second user, referred to as the *follower*, can only see the virtual workspace and verbally communicate with the *leader* regarding its position. Once the position is deemed satisfactory, the *leader* ends the manipulation and is automatically teleported to the newly positioned virtual workspace. Subsequently, the *follower* can use a virtual ray to select the virtual workspace and teleport themselves inside it, thereby recovering the spatial consistency between the users. We have divided the teleportation process into two steps, rather than simultaneously teleporting both users, to prevent unwanted teleportation of the *follower* which can lead to frustration and disorientation.

In the *Co-manipulation* technique, the users manipulate the virtual workspace together to define its future position before teleportation (Figure 3.14-right). Both users use a virtual ray attached to their VR controller to indicate their targeted future positions in the virtual environment. The technique equally incorporates inputs from both users using a physically-based approach. The user-defined targeted positions and the users' future positions in the virtual workspace (represented by their preview avatars) are connected by a mass-spring-damper system (Figure 3.15). This system computes the position and orientation of the virtual workspace, enabling users to manipulate it concurrently. Bending rays [Rie+06] are used to provide continuous feedback on users' mutual actions. The navigation technique switches from individual teleportation to the co-manipulation of the virtual workspace as soon as the two users' targeted future positions are close to each other. Once the users agree on the virtual workspace position, one of them can end the co-manipulation. Both users are then teleported into the newly defined virtual workspace, recovering the spatial consistency between them.

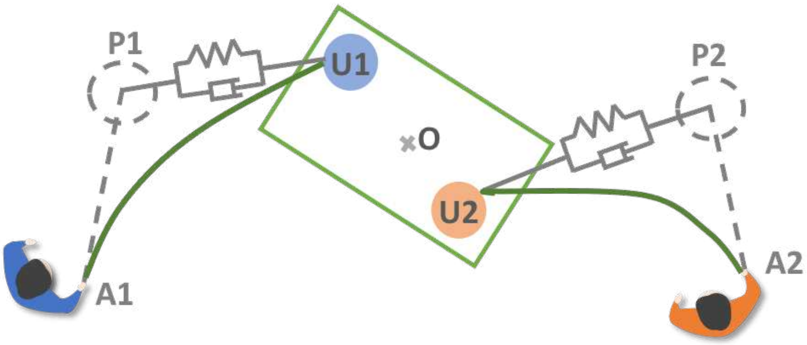


Figure 3.15: *Co-manipulation* technique integrating users' inputs using a physically based system: both a spring and a damper connect the user-defined targeted positions ($P1\&P2$) with the users' future positions in the virtual workspace ($U1\&U2$). Bending rays between the users' VR controller ($A1\&A2$) and their future positions in the workspace ($U1\&U2$) provide feedback on their mutual actions.

We conducted a controlled experiment to compare these two techniques in a virtual riveting task that consisted of individual navigation and collaborative assembly phases. 24 participants were grouped into pairs and equipped with *HTC Vive Pro Eye* VR headsets, while being co-located in a $3m \times 4m$ physical area. Participants first performed individual tasks at separate locations within a virtual factory: one participant prepared the hammer, while the other collected rivets. Next, they regrouped in a designated area to rivet a helicopter shell together (Figure 3.16). To accomplish this, they needed to recover a consistent spatial mapping between themselves. This spatial consistency allowed for direct physical contact between the participants' VR controllers, providing passive haptic feedback as they hammered the rivet. When recovering spatial consistency, the virtual workspace had to enclose three riveting locations: two locations were only known by one participant, while the third one was only known by the other. As a result, they had to negotiate the positioning of the virtual workspace. We hypothesized that *Co-manipulation* would reduce the time spent negotiating the future workspace position and induce better workspace positioning compared with *Leader-and-Follower*.

Results show that *Co-manipulation* significantly reduced participants' time spent positioning the virtual workspace compared to *Leader-and-Follower*, decreasing the overall task completion time. This can be explained by the fact that participants' intents can be communicated through the manipulation with *Co-manipulation*, eliminating the need for verbal descriptions of positioning requirements. Although no significant difference was found between the two conditions regarding riveting time, participants experienced more frequent collisions with their physical workspace boundaries and had to reposition the virtual workspace more often to perform the riveting task with *Leader-and-Follower* than with *Co-manipulation*. This suggests that participants achieve better positioning of the virtual workspace with *Co-manipulation*. However, *Co-manipulation* could introduce conflicts during the manipulation of the virtual workspace. In particular, some participants found it difficult to understand how they influenced the movement of their virtual workspace.

In summary, we have compared two interactive techniques that assist two co-located users in defining the area of the virtual environment where they want to restore a consistent spatial mapping between their physical and virtual positions. The *Leader-and-Follower* technique allows only the *leader* to position the future virtual

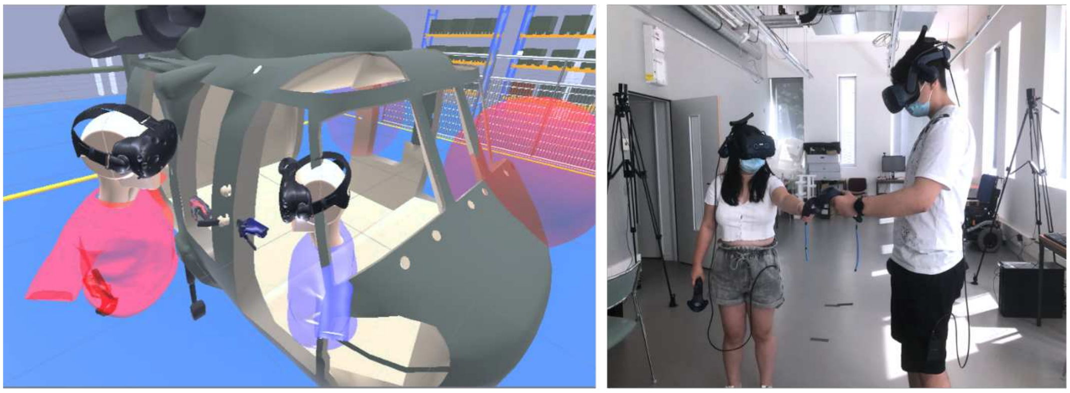


Figure 3.16: (Left) virtual view and (right) real view of the collaborative riveting task on a helicopter shell: users' VR controllers represent a hammer and riveting pliers in the virtual environment. They come into direct contact, providing passive haptic feedback as the rivet is hammered.

workspace, while the *Co-manipulation* technique enables collaborative positioning. Although the *Co-manipulation* technique may be difficult to handle for some users at first, it significantly reduces the time needed to negotiate for the position of the area and enables better placement. Further investigation is necessary to assess these two techniques in various other collaborative scenarios. In particular, the *Co-manipulation* technique can be extended to more than two users, but this may introduce more conflicts and difficulties during concurrent manipulation. In such cases, it could be appropriate to test different spring-damper values, giving users unbalanced control during manipulation and creating an alternative between the co-manipulation and leader-follower approaches.

3.4 CONCLUSION

Appropriate interaction techniques are mandatory to support collaboration in large interactive spaces. In this chapter, I first introduced several interaction paradigms designed to address the specific characteristics of these systems. I especially focused on three key aspects: (i) the large visualization space, (ii) the 3D space available for interaction and (iii) the large physical space surrounding users. All the proposed paradigms support multiple users interacting from different locations within the same interactive space. In a second step, I investigated the impact of these paradigms on co-located collaboration and how they can be extended to provide users with specific features enhancing collaboration. This illustrates examples of how we transition from multi-user interaction to truly collaborative interaction.

For the first aspect, we designed touch-based interaction techniques for creating, manipulating and organizing numerous design alternatives of a 3D object on a wall-sized display. We then studied how pairs of users collaborate during a collaborative design task using these techniques. This study demonstrates that distributing numerous design alternatives on the wall-sized display enhances design exploration and negotiation by increasing the common ground among collaborators.

For the second aspect, we proposed deforming industrial CAD objects in an immersive VR system by physically pulling or pushing on their surface. When multiple users perform such 3D interaction in a shared space, conflicts can arise and they can disturb each other when interacting with the same data. To address this issue, we designed a collaborative AR system that enables users to interact in 3D with distinct versions of the virtual content. Such systems can provide users with the abilities to explore their own design ideas in the 3D space, while also facilitating the sharing of design alternatives with collaborators at a later stage.

For the third aspect, we investigated various techniques for navigating in a virtual environment, taking advantage of the large physical space surrounding users to maximize physical displacements and allow tangible interaction. We then extended the proposed concepts to collaborative navigation with two co-located users. We designed collaborative navigation techniques that enable users to restore a consistent spatial mapping between their physical and virtual positions, after it has been disrupted by individual navigation in the virtual environment. Our findings highlight the benefits of providing appropriate collaborative tools for such tasks rather than relying solely on verbal communication.

The work presented in this chapter mainly targets collaborative design scenarios, including 3D sketching, computer-aided design and industrial assembly tasks. Although the final interaction techniques are specific to each application context, we created more generic concepts which can be extended to various other domains. The ability to generate and distribute a large number of alternatives on a wall-sized display can be applied to many other ideation or data exploration scenarios involving parameter variations. Providing users with both individual and shared virtual content in a 3D space can be useful in many creative applications. Allowing users to navigate individually in a virtual environment while maintaining the capacity to restore spatial consistency between their relative positions can be valuable in any collaborative virtual reality application that involves physical or tangible interaction among users. Moreover, the design processes we employed can be applied in other contexts for customizing the interaction techniques. For example, the prototyping methods used to design interaction on the wall-sized display could be beneficial to adjust the interaction in a wide variety of applications.

This research contributes to the development of novel interaction paradigms for large interactive spaces. However, interacting with such systems is still new to users and not easy to learn and understand. Many challenges remain in standardizing the interaction techniques and making them easier to discover. Given the various types of devices available, ranging from virtual reality headsets to large wall-sized displays, it is crucial to design consistent interaction techniques that allow users to interact seamlessly across the mixed reality continuum [Mil+95; SSW21]. Standardized techniques should avoid users having to relearn the whole set of interaction mechanisms every time they switch devices.

This research also investigates the design of collaborative systems with dedicated collaboration features and demonstrates their benefits for co-located collaboration. However, further evaluations are needed to comprehensively assess these systems given the wide variety and complexity of collaborative scenarios. In addition, future work should consider users with different levels of expertise and explore solutions for adapting interaction to this expertise. Lastly, collaborative systems are now increasingly using hybrid configurations, including both co-located and remote users. Consequently, the proposed collaborative interaction techniques must be extended to support remote users.

4

COLLABORATION AND AWARENESS ACROSS REMOTE SPACES

Large interactive spaces provide new opportunities for remote collaboration as they can connect distant users and create a shared collaborative space. This allows collaborators to interact while being remote and leverage the benefits of each other's interactive systems. Moreover, the large visualization and physical spaces available in such systems offer a wide range of possibilities for enhancing awareness and communication among these users.

Nevertheless, these collaborative environments raise many challenges, including both technical aspects and issues related to interaction and awareness among remote users. Firstly, technical solutions are needed to allow data sharing and collaborative interaction at a distance. It is especially important to handle users with heterogeneous devices and asymmetric setups. Secondly, these systems must also facilitate understanding between users, as distance and technology can alter awareness and communication among them. In particular, large interactive spaces require suitable means of representing remote users, showing their actions and interaction capabilities, as well as transmitting non-verbal communication cues. These solutions should take advantage of these systems to go beyond reproducing the standard face-to-face collaboration that happens when no technology is involved, as proposed by Hollan & Stornetta in their article "Beyond Being There" [HS92].

In this chapter, I present my research on remote collaboration and telepresence systems. The first section focuses on different technical aspects involved when connecting heterogeneous interactive spaces, ranging from wall-sized displays to immersive virtual reality systems. For each aspect, I describe how the technology can be leveraged to support an effective collaboration. The second section explores how video-mediated communication can enhance awareness among remote collaborators. In this section, I detail the design of telepresence systems covering various forms of collaboration, including one-to-one collaboration, one-to-many collaboration or collaboration between users of immersive and non-immersive technologies.

4.1 CONNECTING HETEROGENEOUS SPACES

Remote collaboration across large interactive spaces cannot become widespread if it requires all users to have the exact same physical devices, especially given the wide range of devices currently available. My goal is to design collaborative systems that can accommodate users with heterogeneous devices and asymmetrical setups. In particular, I want to take advantage of the asymmetrical interaction capabilities to foster new collaboration strategies, as presented in our position paper [**Fle+15**b].

This section mainly focuses on the technical aspects of connecting remote users across heterogeneous platforms and enabling communication among them. The first subsection explores data sharing and demonstrates how immersive and non-immersive spaces can be interconnected to support collaboration on computer-aided design (CAD) data in the context of industrial design. This work was published at the *3DCVE workshop* at IEEE VR 2018 [**Oku+18b**] and in a book chapter [**Oku+21**]. The second subsection concentrates on 3D audio for transmitting and rendering remote users' voices. It proposes various spatial audio mappings to connect remote spaces with different sizes and shapes. The related system was presented at the Web Audio Conference 2018 [**Fyf+18**]. Finally, the last section proposes a method for reconstructing live 3D models of users' heads and transmitting them to remote locations. Such models can be used to create and animate realistic avatars of remote users in immersive telepresence or virtual reality systems. This method appeared at Eurographics 2014 [**Fle+14**].

4.1.1 *CAD data synchronization for collaborative modification*

Connecting remote users across heterogeneous platforms can offer several advantages in an industrial design process. It allows multidisciplinary experts to work together despite being located in different branches of a company, but also provides them with specific systems tailored to their needs. For example, style designers may require a large, high-resolution screen to explore and compare multiple alternatives, while ergonomists may prefer an immersive VR system to view the product in context. However, sharing computer-aided design (CAD) data across heterogeneous platforms and modifying it in real time are challenging. While modifying CAD parameters from virtual environments is complex, managing collaborative modifications is even harder, as it requires additional synchronization mechanisms. Our goal was to create a distributed system allowing remote users to modify together native CAD data across heterogeneous platforms.

While distributed systems for collaborative virtual environments have been studied since the 1990s in academic research [**Fle+10c**], only a few studies have addressed collaborative product reviews [Lei+96; LD97] and collaborative VR-CAD applications [Mah+10; AG00]. However, these systems do not support collaborative modification of CAD-part parameters. The *Multi-Agent System* [Mah+10] allows engineers and ergonomists to manipulate the position and orientation of CAD objects across a VR platform and workstations, but it does not allow the shape of the object to be modified through its parameters. *DVDS* [AG00] enables users to create a 3D model with hand gestures in a virtual environment, but it relies on a dedicated CAD system, and does not implement a distributed architecture to share this model across remote platforms. In the previous chapter, I described the design of two interaction techniques that enable non-CAD experts to modify native CAD data in large interactive spaces without using conventional CAD software. *ShapeCompare* (Section 3.1.2) facilitates the generation and visualization of numerous design alternatives on a wall-sized display, while *ShapeGuide* (Section 3.2.1) allows users to deform CAD objects through physical actions in an immersive VR system.

Building upon this work, we created a distributed architecture that synchronizes CAD data across remote platforms and deals with collaborative modification [**Oku+18b**; **Oku+21**]. This architecture is based on an external server, named *VR-CAD server*, which provides centralized access to native CAD data by embedding the CAA API of CATIA V5[1] (Figure 4.1). This server loads and modifies CAD data according to users' requests, using a *labeling* concept [CB04]. The *labeling* maintains a direct link between the 3D geometries displayed in each platform and

---

[1] `https://www.3ds.com/products-services/catia/`

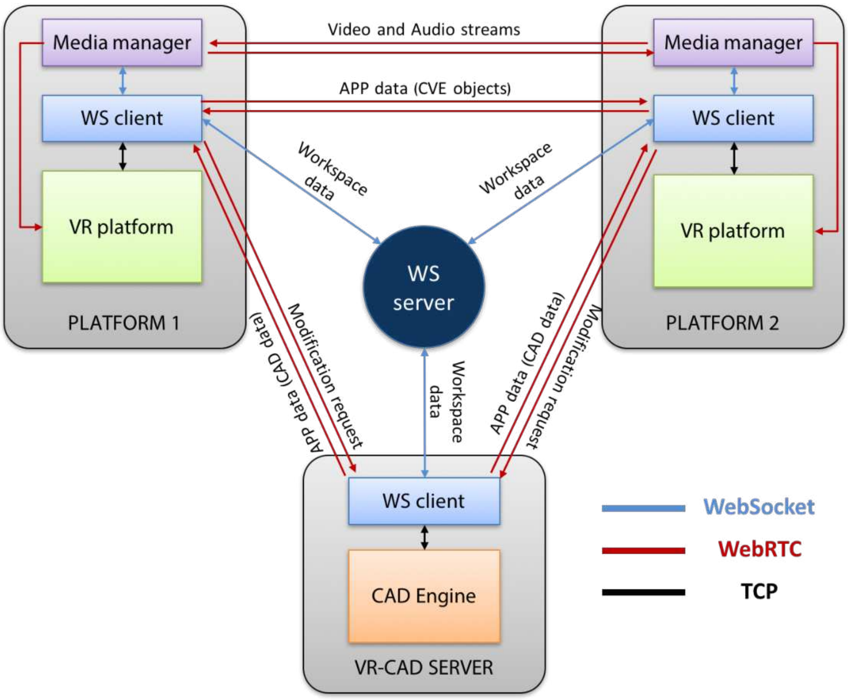


Figure 4.1: Distributed architecture for collaborative CAD data modification across remote platforms: the *VR-CAD server* is responsible for loading, modifying and synchronizing CAD data, while the *WS server* and *WS clients* manage connections between each platform and with the *VR-CAD server*.

the corresponding CAD parameters. Consequently, when users wish to modify a CAD object by interacting with its geometry, the *VR-CAD server* can retrieve the parameters to be modified and generate the desired shape. Once the modification request has been processed, the server sends back the tessellated meshes and a new *labeling* file to each platform, thus updating the visualization.

We used a hybrid network architecture to handle connections between each platform and with the *VR-CAD server*. A centralized architecture connects each platform to the *VR-CAD server* and manages CAD data synchronization. Additional peer-to-peer connections allow fast communication between platforms for all the other data types, including audio and video streams. To implement this architecture, a *Workspace (WS) client* deals with the network communication on each platform and on the *VR-CAD server*. A *WS server* handles authentication and initialization of the connections between all the *WS clients*, but direct peer-to-peer connections are used between *WS clients* to transmit data with the WebRTC[2] protocol. Since the communication layer is independent from the platform technical specifications, this architecture can connect heterogeneous platforms with various visualization systems and interaction devices.

The *VR-CAD server* supports both independent and cooperative modifications of the CAD data. Independent modifications enable several remote users to act on different CAD parameters simultaneously. When users modify multiple parameter values at the same time, the *VR-CAD server* processes the modification requests in the order they are received, and updates the CAD object on all platforms, regardless of the other ongoing modifications. Although this can be a little confusing for users,

---

[2] https://webrtc.org/

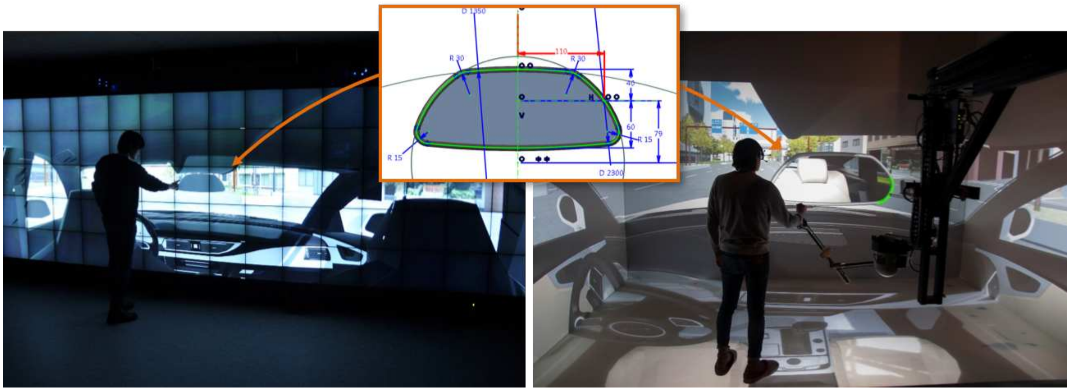


Figure 4.2: Collaborative modification of a car rear-view mirror between (left) a wall-sized display and (right) an immersive VR system. The *VR-CAD server* synchronizes the CAD data of the mirror between both platforms.

it can be effective if they coordinate well and modify complementary parameters. For example, one can modify the radius of a cylinder, while another can change its length. Cooperative modifications allow several remote users to modify the same CAD parameter simultaneously. In this case, the *VR-CAD server* manages concurrent modifications by using a dedicated concurrency control mechanism. In this first prototype, we simply used an averaging technique to combine the parameter values of each user, as proposed by Ruddle et al. [RSJ02]. However, future implementation could explore more sophisticated techniques, such as those studied in previous work [PBF08; ADL10a]. Such cooperative modification can be useful to help remote experts with divergent design constraints negotiate the shape of the CAD object through interaction.

As a proof of concept, we used this distributed architecture to implement collaborative modification of CAD data between a wall-sized display and an immersive VR system (Figure 4.2). We chose these two remote platforms because they have different visualization and interaction features. The wall-sized display has a high-resolution touch screen controlled by a cluster of 10 computers (see description of the Wilder system in Section 2.1.1). The VR system is composed of four large stereoscopic screens controlled by a cluster of 5 computers and a haptic device mounted on a carrier which enables users to interact everywhere in the system (see description of the Eve system in Section 2.1.2). The two platforms are located in remote buildings and connected to separate LAN networks. We set up the architecture by connecting a *WS client* to the master node of each platform and to the *VR-CAD server*. When the master nodes receive CAD data from the *VR-CAD server*, it still has to replicate this data on the slave nodes of the cluster.

We explored a collaborative design scenario where users could benefit from the asymmetric interaction capabilities to collaboratively modify the native CAD data of a car rear-view mirror. A team of style designers could use *ShapeCompare* to quickly generate various alternatives of the rear-view mirror on the wall-sized display, while an ergonomist could use the VR system to sit in the virtual car cockpit and review these alternatives. This enabled the ergonomist to assess live visibility in the rear-view mirror under realistic driving conditions. The ergonomist could also fine-tune the design of an alternative in VR by using the haptic device and the *ShapeGuide* technique. Additionally, we explored a second scenario in which two users perform concurrent modifications of the rear-view mirror within the virtual car cockpit from both the wall-sized display and the VR system (Figure 4.2). Each user was able to modify the mirror shape by pushing or pulling on its surface using the *ShapeGuide* technique. The user in front of the wall-sized display used the finger on the touch screen, while the user in the VR system used the haptic device. When they modify the same CAD parameter, the *VR-CAD server* manages the concurrent modifications as described previously.

In summary, this work mainly focused on the technical aspects of connecting heterogeneous platforms and sharing native CAD among them. We proposed a hybrid distributed architecture that supports collaborative modifications of native CAD data from remote platforms. The CAD data is distributed and synchronized through a dedicated server, while other data and media streams are directly shared between platforms through peer-to-peer connections. We successfully implemented a proof of concept between a wall-sized display and an immersive VR system. Future work should further evaluate the benefits of such an asymmetric system in real collaborative design situations. We should also focus on improving collaborative interaction and providing appropriate feedback of the other users' actions.

4.1.2 *Spatial mapping for 3D audio communication*

When connecting remote users located in large interactive spaces, transmitting voice is a crucial aspect of communication. Using spatialized 3D audio for rendering voices can provide users with additional cues regarding the positions and activities of their remote collaborators. To achieve this, we can map the positions of voice sources to the actual 3D positions of the remote collaborators within their interactive system. However, mapping remote audio spaces with the local 3D space becomes challenging when the interactive systems differ in size and shape. It introduces additional complexities when systems are asymmetric or use different visual representations of the remote collaborators, such as avatars in a virtual environment as opposed to video feeds on 2D displays. Our goal was to create a technical system capable of transmitting audio along with users' 3D positions and rendering spatialized sound, in order to enable us to explore various spatial audio mappings across remote heterogeneous spaces.

Early work investigated spatial audio in telepresence systems [HRB97] and studied binaural audio in such a context [CAK93]. Binaural audio involves recording and rendering distinct sounds for each ear to replicate 3D audio as experienced by users in a real environment. To perceive this binaural audio accurately, users must use headphones. Similarly, Keyrouz and Diepold [KD07] employed binaural audio to allow a teleoperator to perceive the sound of a remote environment in 3D. However, these studies focused on the technical aspects of audio capture and rendering. They also assumed a one-to-one mapping between the recording environment and the rendering space, without exploring alternative mappings.

We created a telepresence system that records users' voices and 3D positions, transmits this data to remote platforms and renders spatialized sound using binaural audio feedback [**Fyf+18**]. In each platform, all users are equipped with wireless microphones and headphones, allowing them to move freely within the system while communicating. Voices are captured through an audio interface and sent to

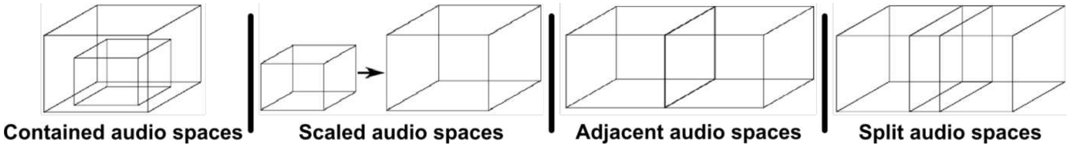


Figure 4.3: Multiple mappings between remote audio spaces of heterogeneous systems.

a media server based on *Kurento*[3]. Users' positions in the 3D space are captured by a *VICON* infrared tracking system. All users wear reflective markers, on their headphones for example. These markers identify each user individually. Audio and 3D positions are transmitted, along with the video, to remote platforms using the WebRTC[4] protocol. The media server receives remote audio streams with their associated 3D positions, and computes the binaural rendering using the *Audiostack*[5] software. It uses both the remote users' positions to compute the voice source locations and the local users' positions to compute the proper binaural audio feedback corresponding to their position and orientation. Finally, *Audiostack* provides users with appropriate audio feedback through their headphones. As a result, users perceive their remote collaborators' voices as coming from specific 3D locations. These locations remain consistent even if the users move or turn their head.

This spatialized audio feedback creates 3D audio spaces that are mapped with the local 3D space of the interactive system, and vice versa in the remote locations. This mapping can be modified or distorted by changing the audio space positions in the local reference frame or by altering the sound quality, with attenuation effects for example. We explored various mapping between audio spaces that are useful to connect heterogeneous spaces with different sizes and shapes (Figure 4.3):

- **Contained audio spaces:** when two remote spaces have different sizes, one obvious solution is that the smaller space is contained within the larger one. The smaller one can be positioned anywhere inside the larger one, allowing a specific placement of the remote collaborators inside the larger space. Although this solution offers real-scale mapping of the two spaces, it can be frustrating for the users in the smaller space to hear others from far away and not be able to join or follow them.

- **Scaled audio spaces:** a second solution when two remote spaces have different sizes is to scale the smaller space to the size of the larger one, and vice versa. This configuration can be useful when the two spaces have the same virtual content displayed on 2D screens of different sizes. The locations of user voice sources are thus consistent with positions relative to the virtual content. However, the speed at which voice sources move may not match the real displacements of the users in the remote location.

- **Adjacent audio spaces:** if we do not want the audio spaces to overlap, they can also be virtually placed adjacent to each other, creating a larger audio space. This configuration works well for telepresence systems with large screens showing the remote locations. It can thus give the feeling that the remote collaborators are "on the other side" of the screens.

---

[3] https://kurento.openvidu.io/

[4] https://webrtc.org/

[5] https://www.aspictechnologies.com/audiostack/

- **Split audio spaces:** a mixed solution consists in overlapping only a subpart of the audio spaces, thus creating different areas within the audio space that may or may not be shared. This configuration is relevant to support different moments in collaboration, and allows users to move around depending on who they want to talk to. For example, users can stay in the shared area to talk with the remote collaborators, then move to the non-shared area when they want to have side discussions with their local collaborators.
- **Distorted audio spaces:** the previous configurations assume that the positions of the voice sources in the local space match the position of the corresponding remote collaborators in the remote space, modulo the possible translation, rotation or scaling required by the configuration. However, the mapping can be further distorted or changed entirely. For example, some systems can display video from remote users in small windows on larger screens, or on mobile devices such as tablets or telepresence robots. In this case, the position of voice sources can be made consistent with the position of the related video streams. Distorted audio spaces open up a wide range of possibilities.

To conclude, we have proposed a technical system that combines audio streaming, motion tracking and spatialized binaural audio in the context of remote collaboration across large interactive spaces. This system allows transmitting users' voices along with their respective 3D positions, and rendering voice sources spatialized within the 3D space of remote locations through binaural feedback. To handle heterogeneous remote platforms, we proposed various mappings between the remote audio spaces and the local 3D space of each platform. This is a preliminary work on how to enhance remote collaboration by customizing these audio mappings. This concept needs to be refined and evaluated in various technical and application contexts. Although the proposed mappings theoretically extend to more than two platforms, we have only tested them in this simple configuration. We must also investigate their impact on collaboration and, especially, how they can support different collaborative dynamics, such as interrupting others' activities, engaging in discussion, initiating side discussions, or transitioning from tightly-coupled to loosely-coupled collaboration.

4.1.3 *3D head reconstruction for immersive telepresence*

Appropriate visual representations of remote users are essential for collaboration across large interactive spaces. They can convey non-verbal cues that are essential for communication, such as eye gaze direction, facial expressions and gestures. However, not all visual representations are suitable for all types of interactive systems when connecting heterogeneous platforms. In particular, video is not well suited to immersive virtual reality systems and 3D displays due to its 2D nature. Avatars can be used in such systems, although facial expressions and eye gaze are usually poorly represented. In this work, we aimed to reconstruct a live 3D model of the users' head to improve avatars and better convey facial expressions and eye gaze to remote collaborators. To easily adapt the proposed system to a wide range of interactive spaces, we targeted a simple solution based on a single consumer level hybrid sensor capturing both color and depth.

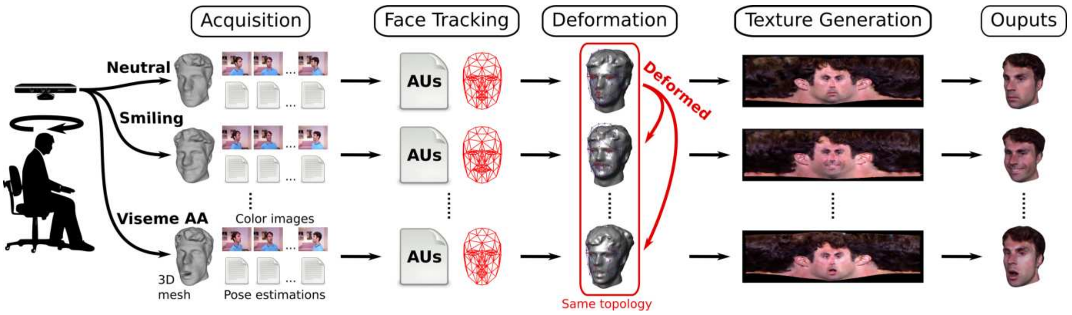


Figure 4.4: Acquisition step of the 3D head reconstruction: data are captured and processed to create a complete and fully textured 3D head model for each facial expression.

3D head reconstruction has been widely studied in the literature, as detailed by Pighin and Lewis [PL06]. Some techniques achieve very accurate 3D models using a large set of high-resolution cameras [Bee+10] or structured lights [Zha+04]. They can also provide an animated version of such head models [Bee+11]. Nevertheless, these techniques require an expensive and complex equipment. At the time we carried out this research, they did also not operate in real time which was not suitable for remote collaboration. Other techniques achieved real-time reconstruction by fitting a deformable face model to the depth data of the users' face captured by a depth sensor [Wei+11; Li+13]. This deformable face model is usually created from a large database of human face scans. As a consequence, it does not fit the specific appearance of the users' head because hair, eyes and interior of the mouth are missing. A colored texture of the face can be generated [Wei+11], but it is static and inconsistencies appear for small face features, including eyes, teeth, tongue or wrinkles. Consequently, these techniques are unable to properly convey facial expressions, which are crucial for non-verbal communication.

We proposed a 3D head reconstruction method that animates the model in real time and makes it suitable for remote collaboration [**Fle+14**]. It uses only a single consumer level hybrid sensor capturing color and depth, such as the Microsoft Kinect used in our implementation. This sensor has to be located in front of the users and does not require any calibration, which makes it easy to install in large interactive spaces. However, this type of sensor provides noisy and incomplete data due to poor sensing quality and occlusions. Our method fuses the noisy and incomplete real-time output of the sensor with a set of high-resolution static textured models captured offline in a preliminary step. The method is decomposed into two steps: an acquisition step that captures and pre-processes data, and a reconstruction step that reconstructs the head model in real time.

For the acquisition step (Figure 4.4), users must spin on a chair in front of the sensor and display different facial expressions during each turn. These expressions include visemes, as well as other variations such as a neutral expression, open mouth, smile and raising or lowering eyebrows. For each facial expression, we use the KinectFusion algorithm [Iza+11] to generate a 3D mesh of the head along with a set of color images captured from different angles around the head. Each image is accompanied by its camera pose estimation relative to the head position. We also use the face tracker from the Microsoft Kinect SDK [Cai+10] to track the head and to characterize the related facial expression in each data set. The tracker provides us with a set of descriptors that are stored with each 3D mesh. After the data

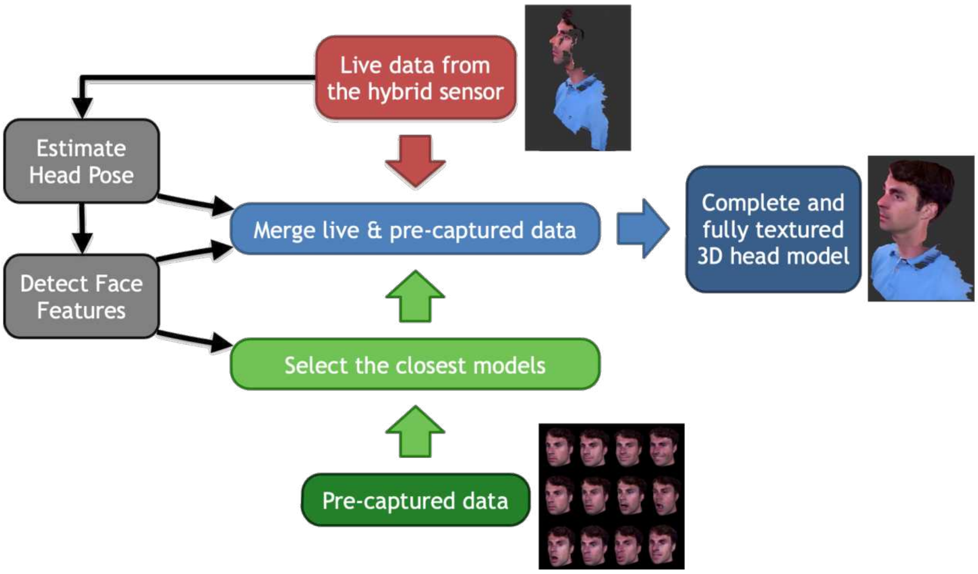


Figure 4.5: Live reconstruction pipeline of the users' 3D head: color and depth data from the sensor are combined with the pre-captured data to create a complete and fully textured 3D model.

capture, we deform the neutral face mesh to fit the other meshes using a method based on cross-parametrization [KS04]. This produces a set of deformed meshes, all with the same topology, which can later be smoothly interpolated to match the current facial expression during the reconstruction step. Finally, a cylindrical texture is generated for each facial expression by combining the color images using a cylindrical projection . We use alignment and smoothing processes to compensate for inaccuracies in camera pose estimations and disparities in color balance and lighting across images. The output of the acquisition step consists of a set of 3D meshes that share the same topology, along with their corresponding cylindrical texture and facial expression descriptors.

For the reconstruction step (Figure 4.5), users need to stand in front of the sensor during the remote collaboration session, as in any videoconferencing systems using a camera. We use again the face tracker from the Microsoft Kinect SDK to estimate head pose and detect users' current facial expression. The descriptors detected by the tracker are then compared with the ones stored with the pre-captured 3D meshes by computing the Euclidean distance. 3D meshes that correspond to the closest facial expressions are selected and interpolated to create a new 3D mesh that matches the current facial expression of the user (Figure 4.6-b). In our implementation, we chose to select the two closest meshes, but it is possible to select more if needed. The cylindrical textures associated with the selected meshes are also interpolated to create a new static cylindrical texture (Figure 4.6-c). This static texture allows us to have a 360° texture of the head with a better resolution than the images directly captured from the sensor. However, the dynamic facial features, such as the eyes, mouth or wrinkles, are not consistent with the users' current face. Therefore, we propose to use the static texture as a background, but to combine it with a dynamic texture extracted from the sensor video stream, which provides the salient features of the face. A gradient is used to extract these features from the dynamic texture (Figure 4.6-d) and, conversely, smooth such features in the static texture. The final texture is obtained by merging these two textures

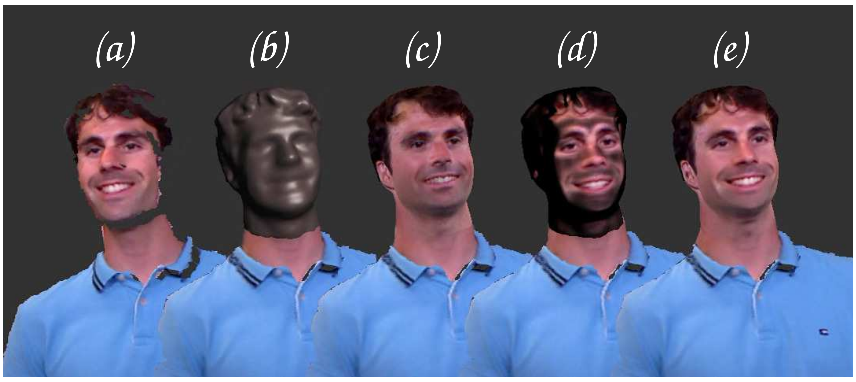


Figure 4.6: Live reconstruction steps: (a) raw data from the sensor and (b) mesh interpolated from the 3D meshes with the closest facial expressions. This new mesh is then textured with (c) the static cylindrical texture, (d) the dynamic texture showing only the salient features, and (d) both textures combined.

(Figure 4.6-e). As a result, a complete and fully textured 3D model of the users' head is reconstructed in real time, accurately preserving facial expressions.

We obtained promising preliminary results showing that our method is able to handle a wide range of subjects and different lighting conditions. Although the 3D mesh was not precise enough to capture small face features, we observed that the dynamic texture can compensate for this limitation. This suggests that integrating elements from live video can be valuable in the context of remote collaboration. Further studies would be required to evaluate our method on a larger scale and better understand which parameters are the most important for communication between remote users. In particular, it would be interesting to assess how the texture quality and 3D mesh accuracy impact the perception of facial expressions. Moreover, our method is well suited to remote collaboration, as the amount of data transmitted over the network is relatively low. Once the acquisition step and system initialization are complete, only the color and depth video streams need to be sent, and the 3D reconstruction can be performed remotely. Finally, our method is highly dependent on the results from the face tracker and most failures occur when it cannot accurately detect facial expressions. This mainly happens when users look away from the camera. However, this could be improved by using a more accurate tracker as our method does not rely solely on this tracker.

In summary, we proposed a method for reconstructing and animating 3D head models of remote users. It relies on a single consumer-level hybrid camera which captures both color and depth. The key features of this approach are an interpolation of pre-captured 3D meshes corresponding to different facial expressions, and a fusion of static and dynamic textures to respectively enhance resolution and incorporate dynamic features extracted from the live video. This work was conducted ten years ago from the writing of this manuscript. Hardware and software solutions have improved dramatically in recent years, especially with the use of machine leaning techniques and the creation of large data sets of human faces. It is now possible to achieve much higher quality results in real time, as proposed by *Pixel Codec Avatars* [Ma+21] for example. However, the closer avatars become to human aspect, the more they raise concerns related to the Uncanny Valley [MMK12]. We must be careful when choosing user representations depending on the application context and further explore the impact of these representations on collaboration.

4.2 ENHANCING AWARENESS WITH VIDEO-MEDIATED COMMUNICATION

Video-mediated communication has long since demonstrated its considerable strengths in remote collaboration systems, as detailed in Section 2.3.1. Video-mediated communication can also be valuable for enhancing users' awareness across remote large interactive spaces, as presented in our position paper [**Fle+15**a]. However, most previous systems are designed for meetings where users sit around a conference table, relying on video as a substitute for face-to-face conversation. These systems do not support large spaces where users move around and work on shared data. Previous work on Media Spaces [BHI93; Mac99] has created systems that support peripheral awareness, chance encounters, locating colleagues and other social activities. Nevertheless, Media Spaces have not explored setups where distributed groups work on shared data in large interactive spaces.

This section explores the design of telepresence systems for large interactive spaces, focusing mainly on non-immersive systems. The first subsection addresses collaboration across wall-sized displays and investigates how to capture and where to display video in such systems. A first perceptual study was published at CHI 2015 [**AFB15**], while the main part of this work was published at CHI 2017 [**Ave+17**]. The second subsection aims to integrate a remote user into a co-located group collaboration, properly conveying gaze direction. These results appeared at INTERACT 2019 [**Le+19**]. Finally, the last subsection details how a laptop or desktop user can collaborate with a remote collaborator wearing an augmented reality headset, taking advantage of multiple video viewpoints on this remote collaborator and the augmented content. This work was published as CSCW 2022 [**FFT22**b] and the related system was demonstrated at IHM 2022 [**FFT22**a].

4.2.1 *Telepresence across wall-sized displays*

Large wall-sized displays are powerful tools for supporting co-located group collaboration, but they can also accommodate remote users by connecting other wall-sized displays. Video-mediated communication is crucial in such remote collaborative scenario to enhance awareness and mutual understanding among users, as previously discussed. Some previous work investigated telepresence systems across two wall-sized displays. Most of these systems aim to display the remote video feed using all the available screen space, creating the illusion of having a glass between the two remote spaces [Wil+10; Dou+12]. However, this does not support collab-

Figure 4.7: Two remote wall-sized displays showing the remote collaborator's video, along with the same content.

oration on shared digital content. Luff et al. [Luf+15] introduced a telepresence system that supports remote collaboration on shared digital objects. This system preserves the physical relations between video and digital objects, allowing users to understand where collaborators are looking or pointing at. Nevertheless, this system relies on a circular configuration of the screens, requires the exact same setups at both locations, and exclusively mimics co-located collaboration without the flexibility to go beyond physical constraints.

Our goal was to design a telepresence system connecting two distant rooms equipped with wall-sized displays showing shared content (Figure 4.7). We explored how this system can combine the shared task space with the shared person space, as defined by Buxton [Bux92]. The former refers to the ongoing task, involving actions such as making changes, annotating and referencing objects. The latter refers to the collective sense of co-presence, involving facial expressions, voice, gaze and body language. Buxton [Bux09] defines the overlap between these two spaces as the reference space, where *"the remote party can use body language to reference the work"*. We first investigated this reference space on a wall-sized display and assessed how accurately users can interpret deictic gestures in a remote video feed. We then explored how to capture and where to display the video feeds on the wall-sized displays by creating a telepresence system based on camera arrays embedded in the display. Finally, we evaluated this system on two collaborative tasks.

4.2.1.1 *Study of deictic gestures*

Referencing shared objects is crucial to support mutual understanding and effective collaboration [Mac99]. Video-mediated communication can affect users' ability to correctly perceive deictic instructions due to technological limitations, including camera and video placements, lens distortion and latency. We focused on a scenario where two wall-sized displays share the same content and simultaneously display a remote user's video feed at the same relative position as the recording camera at the remote location (Figure 4.8). Our objective was to investigate users' ability to determine accurately which shared object the remote user is referencing, without the need for dedicated technology such as telepointers.

While some previous studies have assessed the accuracy of direct eye contact in video-mediated communication [Che02], none of them focused on the accuracy

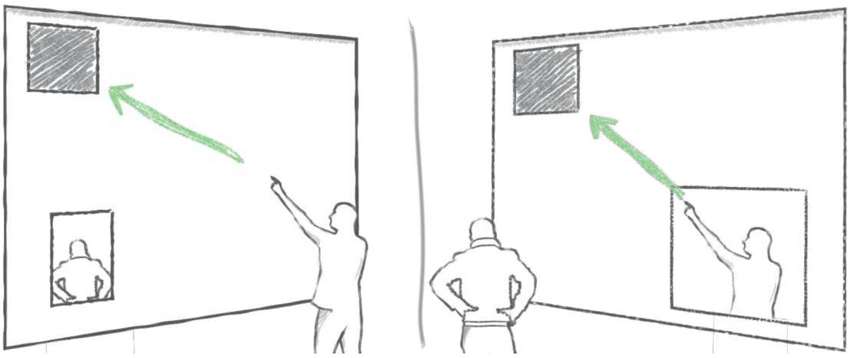


Figure 4.8: Users working on shared objects across two remote wall-sized displays: (left) a user shows a shared object by pointing at it and (right) the remote user tries to understand which object is being pointed through the video.

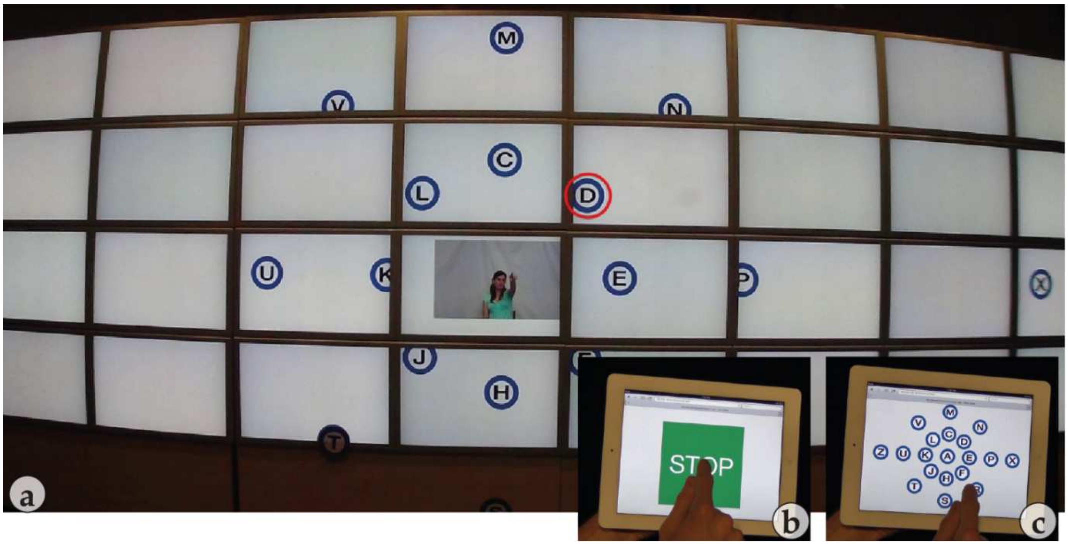


Figure 4.9: Experimental setup from the participants' point of view: (a) the remote user is pointing at target "D" (only highlighted in this figure). To answer, participants first (b) pressed the "STOP" button and then (c) selected the corresponding targets on the tablet.

of remote pointing. Wong & Gutwin [WG10] assessed pointing accuracy, but in a collaborative virtual environment. Users were represented by avatars, which is very different from live video feeds. However, they noted that determining how accurately viewers can interpret pointing direction is a fundamental question to explore before designing support for pointing in remote collaboration systems.

We conducted a controlled experiment to study (i) how accurately participants perceive a reference to a shared object performed by a remote user, either by looking at it or pointing at it with the hand, and (ii) whether the participants' position in front of the wall-sized display influences this accuracy [**AFB15**]. 12 participants looked at a large number of videos of the remote user referencing a specific target on the wall-sized display (Figure 4.9-a). To avoid bias in the experiment, we used pre-recorded videos of three actors playing the role of the remote user. For each video, participants had to indicate on a tablet which target was referenced by the remote user (Figure 4.9-b,c). We controlled three factors:

- 2 techniques used by the remote user to indicate the target: head combines natural head rotation and gaze, while head+arm combines natural head rotation, gaze and pointing with the arm and finger.
- 5 positions of the participants in front of the display: center located in front of the video, farLeft, left, right and farRight respectively located at $2m$ on the left, $1m$ on the left, $1m$ on the right, and $2m$ on the right.
- 19 targets on the wall-sized display, arranged in 8 directions and 3 distances from the central target.

To analyze the results, we decomposed the errors into two measures since the targets were arranged in a circular pattern around the video: distance error and angle error. The unit of distance error is normalized, so that an error of one corresponds to one target closer or further from the center, relative to the designated target. The unit of angle error is in degrees, so that an error of 45° corresponds to one target next to the designated target. First of all, the errors are relatively small, with an overall mean of 0.34 ± 0.52 for distance error and 3.90 ± 12.04 for angle error. This suggests that participants were generally able to accurately identify the referenced target. For the techniques, the distance error is not significantly different when using HEAD or HEAD+ARM, but the angle error is significantly larger when using HEAD+ARM compared to HEAD. Although the effect size is small (1.65°), this result was unexpected. After analyzing the videos, we noticed that the direction of the arm does not always indicate the correct target. In fact, people place the tip of their finger on the line of sight between their eyes and the target, as described in [HC02]. As the video is a 2D representation, it may be hard to perform this 3D geometrical interpretation for viewers, and could lead to errors. For the positions, we observed almost no significant effects of the relative position between the participants and the video on accuracy. This effect on deictic gesture perception is analogous to the *Mona Lisa effect* observed for gaze. The *Mona Lisa effect* describes the fact that the video of a subject looking at the camera is perceived by remote users as looking at them, regardless of their position. At the extreme positions FARLEFT and FARRIGHT, we still measured a slightly higher angle error, but this can be explained by the fact that the observers are looking at video with an angle of 49°, making the task harder.

In conclusion, we assessed how accurately users perceive deictic gestures through video when sharing digital content across remote wall-sized displays. This study shows that users can accurately identify the referenced object, that eye gaze alone can be more accurate than finger pointing, and that the relative position between the viewer and the video has minimal effect on accuracy. Based on these findings, we have derived the following implications for designing future telepresence systems suitable for remote collaboration across wall-sized displays:

1. Additional technical features are not always mandatory to indicate digital objects, as users can accurately interpret gaze and arm pointing. Telepointers and extendable arms [Hig+15] may not necessarily be required if the video is positioned consistently with the content.

2. The arm and gestures are not always needed to indicate digital objects, as users can rely solely on gaze. This allows users to perform deictic actions while holding other interaction tools in their hands.

3. Users can move in front of the wall-sized display, or the video feed can be moved along the display, as the relative position between users and video does not affect accuracy. We can thus consider manipulating the video position on the wall-sized display to meet the requirements of various collaborative tasks.

4.2.1.2 *Design of a telepresence system*

Based on the design recommendations from the previous study, we set out to create a telepresence system supporting video-mediated communication across wall-sized displays. The main challenge was to provide users with audio-video communication as they move in front of the display and interact with shared content.

To determine the optimal camera and video positions, we conducted preliminary observational studies using low-fidelity prototypes. We divided a wall-sized display

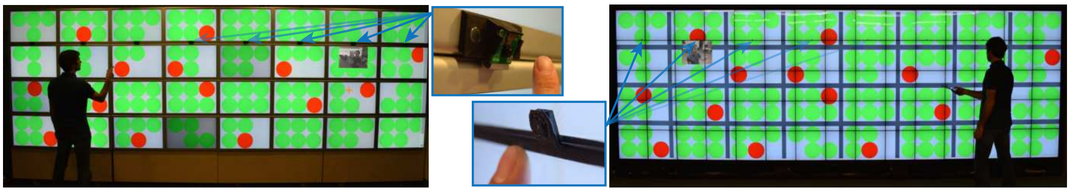


Figure 4.10: *CamRay* provides video-mediated communication between (left) WILD and (right) WILDER wall-sized displays. (Center) close-ups on the cameras embedded in the displays.

with a curtain to simulate two remote locations and simplify the technical setup. The first prototype included two tablets running videoconferencing software, each held by a helper to test multiple placements. Two participants had to create a slideshow presentation from their respective location. We noticed that they looked at the content much more than the video feed. In fact, they only looked at each other when they disagreed or needed to discuss a specific issue. After the debriefing with participants, we hypothesized that they might have looked at the video more often if it did not require switching screens and decided to place the video on the wall-sized display.

The second prototype used two cameras at each location: a front-facing camera attached to the screen and a back camera located at the back of the room, facing the screen. At each location, three video feeds were also displayed: on the left, a window displayed the remote front-facing video with a small thumbnail showing the local front-facing video, and on the right, another window displayed the remote back-facing video. Two participants had to sort research papers to prepare the related work section of a publication. Papers were laid out on the large screen at each location, with their position and current page synchronized. We observed that participants physically moved to a specific video window depending on the task at hand. They used the front-facing video to discuss paper content or how to cluster papers, while they used the back-facing video to understand which paper the other was refereeing or where the other was pointing. However, they had to interrupt their work to glance at video windows, which was perceived as annoying. We concluded that we should be able to capture users' faces even as they moved along the screen and to display the video feeds in a flexible manner. We identified two requirements for video placement, each corresponding to specific phases of the collaboration: one should support face-to-face conversations, while the other should support the use of deictic instructions.

To meet these requirements, we created *CamRay* [**Ave+17**], a telepresence system connecting remote wall-sized displays. We implemented a prototype of this system between the WILD and WILDER platforms, located in two different buildings (see descriptions of the two systems in Section 2.1.1). *CamRay* uses an array of eight cameras embedded in each display, capturing the users' faces (Figure 4.10). The cameras are equally spaced along the horizontal axis of both displays and located on the nearest bezel above users' eye level. We used *Raspberry Pi* camera modules, each one connected through a ribbon cable to a *Raspberry Pi*[6] located at the back of the display (Figure 4.11). Each *Raspberry Pi* captures video with a resolution of 800 × 600 pixels, encodes it in H.264 and streams it to a dedicated computer over UDP using

---

[6] https://www.raspberrypi.org/

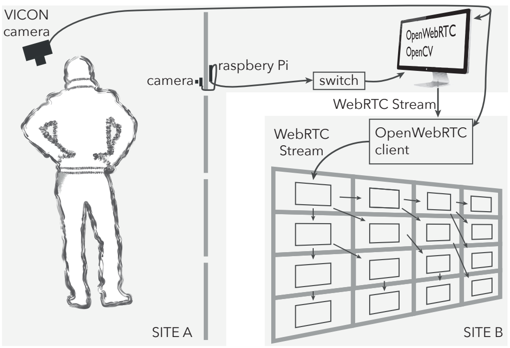


Figure 4.11: *CamRay* architecture transmitting video from site A to site B. A similar setup is used to transmit video from site B to site A.

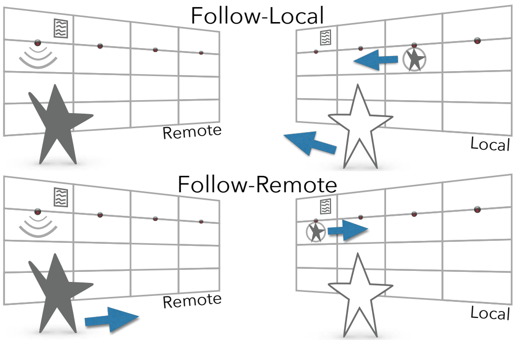


Figure 4.12: The two video modes of *CamRay*. The arrows show the participant whose position controls the video at the local site.

GStreamer[7]. Users' position is tracked by a *VICON* infrared tracking system and also sent to the same computer. This computer executes a C++ application based on *OpenCV*[8] to select the video feed of the camera in front of the user using tracking data. Finally, the application streams the selected video along with the tracking data to the remote wall-sized display using the WebRTC[9] protocol.

On the remote location, a server receives the WebRTC video stream along with the tracking data, and transmits it to the visualization cluster that controls the wall-sized display. Each node of the cluster runs a web application based on *NW.js*[10]. This application is able to display the WebRTC video stream and to forward it to other nodes. Only a specific node in the cluster receives the stream from the server, and forwards it to 2 or 3 other nodes, which in turn forward it to two or three other nodes, and so on (Figure 4.11). This tree pattern allows us to transmit the video stream to all cluster nodes with low latency and avoid overloading the server with multiple video streams. As a result, a video window showing a remote user can be displayed, spanning several screens if required, and moved all over the display. This window can appear on top of the content displayed on the wall-sized display. In addition, the server receives both the positions of the local and remote users, and can use this information to define the placement of the video window.

Based on the observational studies, we implemented two modes for positioning the video window on the wall-sized displays using *CamRay* (Figure 4.12):

- With *Follow-Local*, the video window follows the horizontal position of the local user, providing constant visual contact with the remote collaborator. This mode creates a virtual face-to-face, where both remote users are always visible to each other even when located at different positions in front of their respective displays.

- With *Follow-Remote*, the video window follows the horizontal position of the remote user, conveying his relative position to the shared content. This mode allows users to accurately interpret deictic instructions made by the remote

---

[7] https://gstreamer.freedesktop.org/
[8] https://opencv.org/
[9] https://webrtc.org/
[10] https://nwjs.io/

collaborator as the video window has a consistent position, according to the shared content.

In both modes, the video window does not move continuously, but is snapped under each camera of the array. The video is thus placed congruently with a camera, allowing direct eye contact between users. In addition, the video is horizontally mirrored to maintain a spatial consistency between the video and the shared content: a user looking to the left is therefore displayed as looking to the left in the video at the remote location. Consequently, the remote user is seen as standing behind the display, as in Clearboard [IK92]. Moreover, we do not display any feedback of the users' own video, as nobody used it during our observations. Some participants even reported that they trusted the system to capture them properly, since they were not responsible for adjusting the camera position.

The two proposed modes support different aspects of non-verbal communication, including eye gaze, facial expressions and gestures. In particular, Fussel et al. [Fus+04] distinguish two categories of gestures in video-mediated communication: *"pointing gestures, which are used to refer to task objects and locations, and representational gestures, which are used to represent the form of task objects and the nature of actions to be used with those objects"*. We hypothesized that each method is best suited to support different types of non-verbal cues: *Follow-Remote* consistently positions the video relative to the remote users' position and content, facilitating the accurate understanding of pointing gestures, while *Follow-Local* maintains constant visual contact, making it easier to perceive eye contact, facial expressions and representational gestures. To test this hypothesis, we ran two controlled experiments comparing both modes on two collaborative tasks. The first experiment studied *CamRay* during a data manipulating task that relies on pointing gestures. The second experiment assessed *CamRay* in a knowledge-sharing task that benefits from easy perception of eye contact, facial expressions and representational gestures.

4.2.1.3 *Evaluation on a data manipulation task*

In a first controlled experiment [**Ave+17**], we aimed to assess the ability of *CamRay* to properly convey pointing gestures between two remote wall-sized displays. To achieve this, we needed a data manipulation task that requires the production and interpretation of such gestures. We drew inspiration from the disc classification task designed by Liu et al. [Liu+14]. In a co-located collaborative situation [Liu+16], they explored a condition in which one participant instructed another on where to classify discs on a wall-sized display. They observed that this condition mainly relied on deictic instructions. We thus implemented a remote version of this condition. In this version, an *Instructor* had to determine how to classify discs and give instructions to a remote *Performer* who performed the manipulation (Figure 4.10).

Both wall-sized displays were divided into 32 containers, each capable of holding up to 6 discs. On the *Instructor*'s wall-sized display, discs were labeled with small letters which indicated how to group discs in containers. All the discs in the same container needed to have the same letter to be considered properly classified. Properly classified discs were highlighted in green, while misclassified discs were shown in red. The *Instructor* was not able to move discs. On the *Performer*'s wall-sized display, red and green discs were displayed without labels. The *Performer* was able to move discs with a pointing device.

12 pairs of participants performed the classification task with three video conditions: *Follow-Local*, *Follow-Remote* and a control condition, named *Side-by-side*. This control condition used a fixed video window on a separate screen on the left side, perpendicular to the wall-sized display. Each participant alternately assumed the roles of *Instructor* and *Performer* for each video condition, completing the task twice in each role. The experimental setup was composed of the WILD and WILDER wall-sized displays (see system descriptions in Section 2.1.1). We hypothesized that participants would perform the task faster and rely more deictic instructions with *Follow-Remote* than with *Follow-Local* and *Side-by-side*.

The main results demonstrate that participants classified discs significantly faster with *Follow-Remote* than with *Follow-Local* and *Side-by-side*. There are three reasons for this increase in performance. First, *Performers* followed the *Instructors*' position and gaze more closely, and were faster to drop the disc at the correct spot. In particular, the results revealed that the distances between the *Performers*' cursor and the *Instructors*' position, as well as between the *Performers*' cursor and the *Instructors*' estimated gaze point, were smaller with *Follow-Remote* than with the other conditions. Second, participants used more deictic instructions and fewer words with *Follow-Remote* than with the other conditions, reducing the time spent by the *Instructors* giving verbal instructions. Third, participants made fewer misunderstanding errors with *Follow-Remote* than with the other conditions, also reducing the time spent correcting errors. In addition to these results, qualitative feedback showed that a large majority of participants preferred *Follow-Remote* when playing the role of *Performers* (22/24), while *Side-by-side* was ranked first twice. Surprisingly, only half the participants preferred *Follow-Remote* when playing the role of *Instructors*, while *Follow-Local* was ranked first ten times and *Side-by-side* was ranked first twice. This preference might be due to the fact that *Instructors* liked seeing their remote collaborator's face as they gave instructions to check for understanding.

In summary, *Follow-Remote* proposes to display video consistently to the shared content, according to the remote user's position. The results demonstrate that participants were better able to understand deictic instructions with *Follow-Remote* than with the other conditions, reducing the overall cost of communication, as explained by Clark and Brennan [CB91]. As a consequence, *Follow-Remote* provides better performance on the data manipulation task. Nevertheless, some participants preferred the constant visual contact created by *Follow-Local* when checking for their collaborator's understanding. These potential benefits of *Follow-Local* should be further explored in tasks that involve more discussion and knowledge-sharing.

4.2.1.4 *Evaluation on a knowledge-sharing task*

While the first experiment focused on deictic instructions, this second experiment aimed to explore how *CamRay* could convey representational gestures, along with eye contact and facial expressions. We believed that the persistent face-to-face provided by *Follow-Local*, even when users move in front of the display, could be valuable for better perceiving these non-verbal communication cues. Our goal was to design a task involving discussion and knowledge-sharing that would benefit from these specific cues. We drew inspiration from a realistic scenario in which two experts have to combine their knowledge to resolve a problem.

We created a task in which an *Instructor* sees an image located at a random position on the wall-sized display and has to describe it to a remote *Performer*. At

Figure 4.13: (Left) sign language images and (right) Khmer characters used in the knowledge-sharing task to evaluate *CamRay*.

the remote location, the *Performer* has to search for this image among a set of 21 images spread all over the wall-sized display. In the task, the two wall-sized displays do not show the exact same content. This task mimics the scenario in which an expert retrieves and shares some knowledge with a collaborator, who has to find this information in a large data set. By using only images, the task eliminates the need for personal judgments or negotiation when choosing among possibilities. It also does not require participants to memorize or process information, thus mitigating potential biases in the experiment. To ensure that participants perform representational gestures, we conducted pilot tests with various types of images. We selected images of sign language and Khmer (Cambodian) characters (Figure 4.13). The sign language images are straightforward to describe because participants can replicate the hand poses. The Khmer characters are more difficult to describe and require a combination of gestures and speech.

6 pairs of participants performed this task with the same video conditions as in the first experiment: *Follow-Local*, *Follow-Remote* and *Side-by-side*. For each video condition, participants swapped roles and completed the task twice in each role: once with sign language images and once with Khmer characters. The experimental setup was also composed of the WILD and WILDER platforms. We hypothesized that participants would reach better performance and produce of more representational gestures with *Follow-Local* than with *Follow-Remote* and *Side-by-side*.

The overall results did not reveal strong effects of the video conditions in terms of performance, including task completion time and errors. We believe this is due to the fact that participants often decided to walk towards the video in *Follow-Remote* or *Side-by-side*, thus recreating the face-to-face condition. The results demonstrate that participants traveled longer distance with *Follow-Remote* and *Side-by-side* than with *Follow-Local*. They also synchronized more often their relative position with *Follow-Remote*. As a consequence, it is difficult to measure difference in terms of performance, as participants could potentially benefit from face-to-face conversation in all video conditions. We still noticed that participants made fewer errors with *Follow-Local* than with *Follow-Remote* for the sign language images. This difference could be explained by the fact that *Instructors* moved away from the described image in *Follow-Remote*, and sometimes forgot the exact hand gestures to perform, as describing this type of image relies heavily on representational gestures. Concerning the production of representational gestures, we did not observe significant differences overall. However, we were surprised to notice that *Instructors* used significantly more gestures with *Follow-Local* than with the other conditions when replying to clarification requests. We hypothesized that participants spontaneously moved to create a face-to-face situation when they provided the initial explanation. But, when a clarification was required, participants were not always facing each other with *Follow-Remote* and *Side-by-side*. Finally, qualitative feedback shows that participants preferred *Follow-Local* for this task over the other two conditions, regardless of whether they were *Instructor* or *Performer*.

While the experiment does not reveal strong evidence that *Follow-Local* improves performance in a knowledge-sharing task, it does provide some hints that this condition encourages the production of representational gestures and makes their interpretation more accurate. The results also show that the technological hindrance is less pronounced with this condition. In particular, participants are not required to synchronize their position and need to travel less. As a result, participants preferred the *Follow-Local* condition for such a knowledge-sharing task and perceived a lower task load compared to the *Follow-Remote* condition.

4.2.1.5 *Summary*

In this work, we demonstrated in a first experiment that video feed displayed on a wall-sized display can properly convey deictic instructions, including pointing and gazing. Based on this observation, we designed *CamRay* to support video-mediated communicate across wall-sized displays. It embeds an array of cameras on each display to capture users as they move in the interactive space. We also proposed two methods for positioning the video feed on wall-sized displays according to local and remote users' positions: *Follow-Local* creates a virtual face-to-face, while *Follow-Remote* positions the video consistently relative to the shared content.

Two controlled experiments show that each method has its own advantages, making them suitable for different collaborative tasks. *Follow-Remote* supports deictic instructions for a data manipulation task, while *Follow-Local* supports representational gestures for a knowledge-sharing task. Nevertheless, these results are not clear-cut for *Follow-Local* and its potential benefits should be further studied, taking into consideration other non-verbal cues, such as eye contact and facial expressions. In particular, this method could be valuable for discussion and negotiation tasks. However, operationalizing such tasks in a controlled experiment is challenging, and the evaluation of collaboration should be extended beyond simple performance metrics. Given the advantages of both methods, future work should also explore how to seamlessly integrate them without hindering the collaboration process or overloading users.

Although the first prototype was implemented for two remote users in separate wall-sized displays, *CamRay* can accommodate more than one user per location, as the tracking system can individually identify multiple users. It can also scale to more than two locations, as the server can receive multiple WebRTC connections simultaneously. However, further developments would probably be necessary to support large groups at one location or numerous remote locations. In terms of collaboration, we need to explore further the collaborative behaviors that arise in such large groups, including coupling styles and territorial dynamics.

4.2.2 *Perception of a remote user's gaze direction*

While large interaction spaces can foster collaboration within a co-located group, integrating a remote user into such collaboration remains a challenge for current telepresence systems. In this work, we considered a simple scenario in which *co-located collaborators* sit around a table containing various physical artifacts, such as paper printouts or 3D mock-ups. To integrate a remote user in this scenario, most current telepresence systems use a screen and a camera situated at one edge of the table to support video-mediated communication. The wide perspective provided by this side camera makes it difficult for the remote user to see the physical artifacts on the table. In contrast, the co-located collaborators have a closer view of the remote user, with a much narrower perspective. The difference in perspective, along with the offset between the camera position and the video position at the remote location, does not allow co-located collaborators to properly interpret the gaze direction of the remote user. This hinders communication as co-located collaborators can easily understand each other's gaze direction, but struggle to do so for the remote user, potentially excluding this user from the collaboration. Our goal was to create a telepresence system that accurately convey the remote user's gaze direction to the group. The co-located collaborators should be able to understand if the remote user is gazing at one of them or at specific physical artifacts on the table.

Gaze is crucial for the collaboration, as it helps predict conversational attention [Ver99; Ver+01], perceive references to physical objects [ATI18], support remote instructions [HYS16; Yao+18] and enhance users' confidence [Akk+16]. Failing to properly convey gaze can lead to confused communication [VVV00], reduced effectiveness [MG02] and extra efforts to accomplish collaborative tasks [Akk+16; HYS16]. Previous work has explored gaze perception in remote collaboration, but has mainly focused on conveying either gaze awareness between distant users [SBA92; NC05; Gig+14] or gaze on shared digital content [IK92; KK06], leaving the problem of gaze awareness towards physical artifacts under studied. In addition, such systems often require specialized and complex hardware setups on the remote user's side [PS14; Ots16; Got+18], which might be unrealistic for traveling users.

We created *GazeLens* [**Le+19**], a video conferencing system designed to improve co-located collaborators' ability to interpret the remote user's gaze (Figure 4.14). At the group location, a 360° camera is located at the center of the table, and captures a panoramic video of the collaborators seated around it. A ceiling-mounted camera

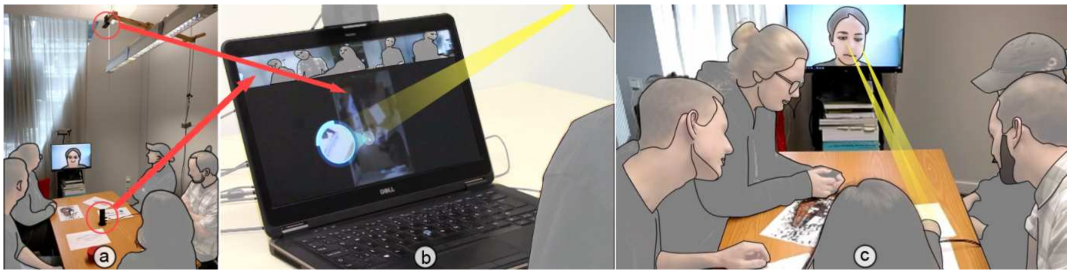


Figure 4.14: *GazeLens* system: (a) a 360° camera and a ceiling-mounted camera respectively capture the co-located collaborators and the physical artifacts; (b) the video from the two cameras are displayed on the remote user's screen, with a virtual lens guiding attention towards a specific screen area; (c) the remote user's gaze is properly aligned towards the observed artifact at the group location.

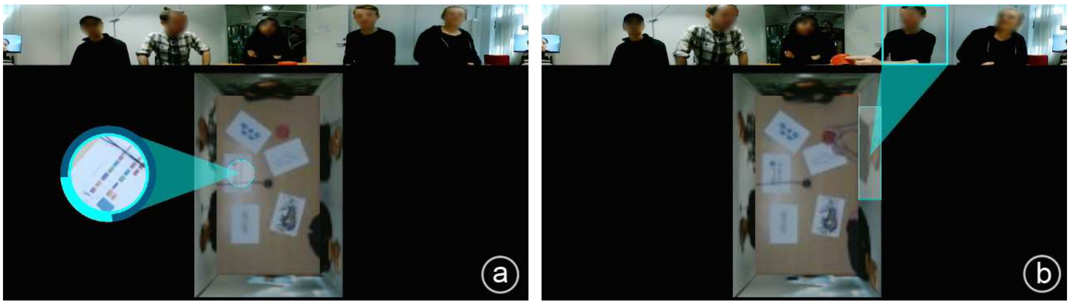


Figure 4.15: *GazeLens* interface: (a) a magnifying lens shows a close-up view of a physical artifact, and (b) a lens indicates a collaborator's position around the table.

captures the physical artifacts on the table, minimizing occlusions. At the remote location, the user has a standard computer equipped with a webcam on top of the screen. *GazeLens* combines the video feeds of the two cameras in a unified interface (Figure 4.15). We designed this interface to strategically direct the remote user's attention toward specific screen areas, allowing co-located users to more accurately interpret the remote user's gaze direction. The 360° video is displayed at the top of the screen, just under the webcam, reproducing eye contact for co-located users when the remote user looks at them in the video feed, as suggested by Chen [Che02]. The top-view video is displayed in the middle of the screen with the correct orientation and aspect ratio relative to the actual table. We used a focus-based approach that mimics foveal and peripheral vision to maximize variation of the remote user's gaze. The top-view video of the table appears slightly blurred, and the remote user can use a magnifying lens to obtain a sharper and closer view of the physical artifacts. This lens is positioned at a specific location on the screen, guiding the remote user's gaze in the right direction with respect to the webcam position (Figure 4.15-a). As physical artifacts can have various orientations on the table, we provide a rotation tool on the lens to rotate its content if needed. Finally, to keep the remote user aware of the co-located collaborators' position around the table, a second view of the 360° video is wrapped around the table top-view video in the interface. This additional view is slightly blurred, and a square lens connects it to the 360° video displayed at the top of the screen, helping the user in correlating these two views (Figure 4.15-b).

We conducted a first controlled experiment to evaluate the effectiveness of *GazeLens* in conveying the remote user's gaze in comparison to a conventional videoconferencing system (Figure 4.16). This baseline used a wide-angle camera to capture the entire room at the group location and displayed the corresponding video in full-screen mode on the remote computer, instead of the *GazeLens* interface. To minimize experiment bias, three actors assumed the role of the remote user. We recorded multiple videos of these actors looking at 14 targets under the two video conditions. 9 targets were arranged in a 3×3 grid on the table, while 5 targets were located at the co-located collaborators' position around the table (Figure 4.16-a). 12 participants took the role of a group member and looked at the pre-recorded videos while sitting at two different locations around the table: in front and on the side of the screen where the remote user's video is displayed (positions C and A in the figure). After viewing each video, they were asked to indicate which target the actor was looking at. We hypothesized that *GazeLens* would improve accuracy of gaze interpretation for both sitting positions compared with the baseline.

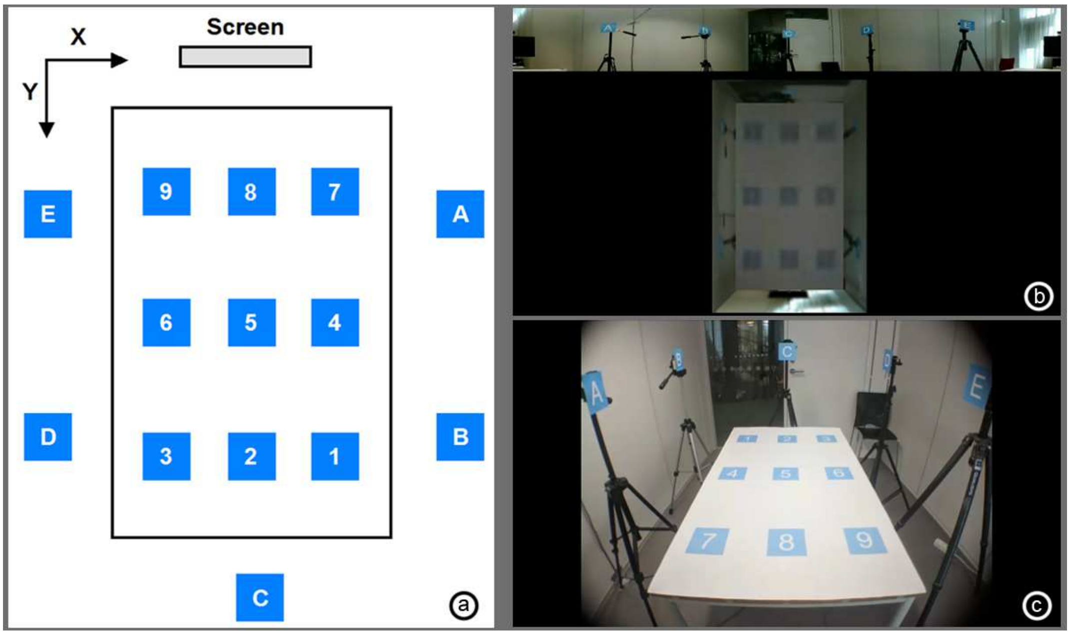


Figure 4.16: Experimental setup of the first study. (a) Multiple targets were laid out at the group location. These targets were viewed at the remote location through the (b) *GazeLens* interface or (c) a conventional videoconferencing system using a wide-angle camera.

Results show that *GazeLens* significantly increases the overall gaze interpretation accuracy of the participants compared to the conventional videoconferencing system. The position of the participants with respect to the screen does not influence these findings. In addition, participants can easily distinguish whether the remote user is looking at a co-located collaborator or at a physical artifact on the table, with over 85% accuracy. However, the results are less good when it comes to distinguishing artifacts on the table, although they were still better than with the conventional videoconferencing system. As a consequence, we decided to conduct a second experiment focusing on the physical artifacts.

A second controlled experiment used the same experimental setup. The only distinction was that the targets were arranged on the table exclusively, and with two densities: 9 targets in a 3×3 grid or 25 targets in a 5×5 grid. 12 participants took part in this experiment. We hypothesized that *GazeLens* would improve accuracy of gaze interpretation for both target densities compared with the baseline.

Results show that *GazeLens* significantly improves gaze interpretation accuracy for table artifacts for sparse, but also dense arrangements, compared with the conventional videoconferencing system. When using *GazeLens*, the accuracy reaches approximately 54% with the sparse layout versus 26% with the baseline. With the dense layout, it drops to around 26% for *GazeLens* and 12% for the baseline. We also analyzed lateral and depth errors. Although *GazeLens* outperforms the conventional system for both types of errors, it results in more depth errors compared to lateral errors. This could be explained by the fact that vertical screen space is limited in our interface, but also by the fact that vertical gaze direction is harder to interpret compared to the horizontal one, as studied by Chen [Che02].

As the two previous experiments focused on the co-located collaborators' perception, we also conducted a preliminary study to gather feedback from the remote user's perspective. Five pairs of participants performed a puzzle-solving task under the two video conditions previously described. In this task, the remote user instructed a group member on how to arrange physical puzzle pieces on the table to match a predefined pattern. Participant swapped roles, taking turns as the remote user and the group member. We gathered qualitative feedback through interviews at the end of the experiment. Only one of the ten participants reported difficulties when using the *GazeLens* interface. He would have preferred another solution to activate the lens than a mouse click. Apart from that, all participants mentioned that it was easier to see the puzzle pieces with the *GazeLens* interface and preferred this condition over the conventional videoconferencing system.

Overall, *GazeLens* provides a telepresence interface that guides a remote user's gaze, enhancing gaze interpretation in video-mediated communication with multiple collaborators located in the same meeting room. Controlled experiments demonstrate that *GazeLens* improves the ability of these co-located collaborators to distinguish whether the remote user is looking at them or at physical artifacts on the meeting room table. *GazeLens* also enhances their accuracy in determining which specific physical artifacts on the table are referenced by the remote user, although there is still room for improvement. In particular, we designed *GazeLens* to be simple enough to be deployed in any meeting room with any camera, but we did not consider camera position, camera focal length, screen size or distance between the screen and the table when defining the lens position. Although achieving geometrically corrected gaze in video-mediated communication is almost impossible, our system could be improved by integrating these parameters. However, this could require some configuration and calibration steps, which can be time-consuming. It would be interesting to explore various trade-offs and find out which parameters have the biggest impact on the accuracy of gaze direction interpretation. Finally, future work should also explore how *GazeLens* can be extended to support multiple remote users. Each remote user can be represented by a dedicated screen around the table, but it could be valuable to investigate solutions using a single large screen, as most meeting rooms are usually equipped with only one screen dedicated to video-meditated communication.

4.2.3 *Exploration of a remote augmented reality workspace*

Augmented reality (AR) makes it possible to create large interactive spaces by integrating virtual content in any physical space. Nevertheless, sharing this virtual content with a remote collaborator is a challenge, especially when this user does not have access to AR or VR equipment. Yet, such asymmetrical collaboration configurations are common today in many circumstances, as more and more collaborators travel or work from home. While video-mediated communication plays a crucial role in enhancing remote collaboration in such contexts, it cannot provide a comprehensive understanding of the AR workspace including both physical and virtual content. Our objective was to establish an effective collaboration between augmented reality and remote desktop users, leveraging the benefits of AR for the remote user.

Several AR technologies propose to video-stream the AR user's perspective. However, these solutions do not provide a view of the user, thus failing to convey non-verbal cues such as gestures, body postures, or facial expressions. Moreover,

Figure 4.17: Remote-view configurations for the first study: (left) HEADSET VIEW, (middle) EXTERNAL VIEW and (right) VIRTUAL VIEW. Remote participants give instructions to the AR user on how to arrange 3D shapes on a virtual support.

the remote user's viewpoint is limited to the AR user's perspective, which hinders the remote user's ability to adequately perceive and explore the AR workspace. Nevertheless, ensuring view independence enhances collaboration performance, as reported by Tait and Billinghurst [TB15]. Some systems create a 3D reconstruction of the AR workspace, allowing the remote user to navigate independently in the 3D scene. However, this 3D reconstruction requites heavyweight and complex hardware setups [AAT13; Bai+20], consumes large communication bandwidth, is affected by network outages [Ahs+21] or imposes significant constraints on the view possibilities of the remote user [Gau+14; Moh+20]. In addition, the AR user is often reconstructed in 3D along with the environment, resulting in a poor user representation that reduces expressiveness and creates an "*uncanny valley of XR [extended reality] telepresence*" [Jon+21].

In this work, we targeted an asymmetric collaboration scenario between a local *AR user* wearing an optical see-through headset (Microsoft Hololens 2) and a *remote user* who participates from a distance through a desktop application. We restricted our design space to lightweight setups using only a single external depth camera on the AR user's side, in addition to the camera of the headset. This external camera could be easily replaced by a webcam and a smartphone as many devices are now equipped with depth sensors. The remote user simply uses a standard laptop or desktop computer with a webcam. We aimed to enhance video-mediated communication between these two users with new visual and interaction modalities.

As a first step, we conducted a user study [**FFT22b**] to investigate the trade-offs associated with different AR workspace representations and scene viewpoints. 24 participants took the role of the remote user and instructed an experimenter, who acted as a confederate, to accomplish a puzzle-solving task in AR. Participants were presented with a randomly generated pattern composed of 8 puzzle pieces among the 18 available in the AR workspace. The experimenter's task was to replicate this pattern on a virtual plane placed on his table. Participants provided instructions under three conditions (Figure 4.17):

- HEADSET VIEW: participants viewed an augmented video from a first-person viewpoint provided by the AR headset camera.
- EXTERNAL VIEW: participants viewed an augmented video from a third-person viewpoint provided by the external camera.
- VIRTUAL VIEW: participants viewed a fully virtual representation of the 3D scene. They could freely navigate in the scene and choose their own viewpoint. No information regarding the physical environment was visible, but a simplified avatar represented the AR user.

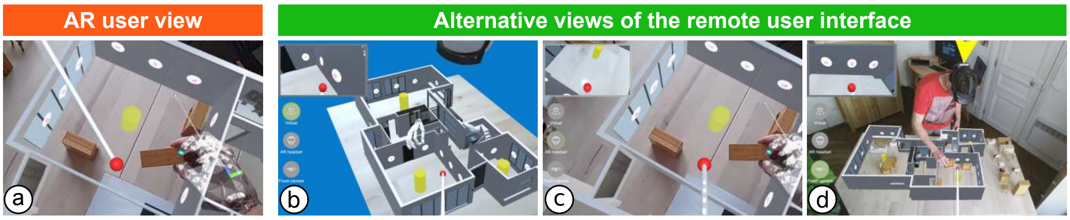


Figure 4.18: A remote user guides an AR user achieving a physical furniture arrangement task in a virtual 3D house model with *ARgus*. (a) The AR user view displayed in the headset and the three AR workspace representations combined in the desktop interface of the remote user: (b) a fully virtual view, (c) an augmented first-person view, and (d) an augmented third-person view.

After completing the task, participants rated the perceived difficulty for different components of the task and their overall preferences. The results indicate that each view configuration has its own qualities that are difficult to substitute using the other views. The External View provides a global perception of the AR workspace and helps participants search for puzzle pieces. The Virtual View supports independent navigation, helping participants give instructions from a convenient and stable viewpoint. Finally, the Headset View is effective for perceiving the AR user's actions and communicating egocentric instructions.

Building on these findings, we focused on combining these multiple representations and providing remote users with direct control over their use. We designed *ARgus* [FFT22b; FFT22a], a multi-view video-mediated communication system that combines the three representations through interactive tools for navigation, previewing, pointing, and annotation (Figure 4.18). *ARgus* receives the augmented video from both the AR headset and the external depth camera. It also maintains a synchronized version of the virtual scene, and can generate virtual views from any location. This enables the remote user to seamlessly switch between the Headset View, the External View and any viewpoint of the Virtual View. Additionally, *ARgus* offers the ability to display live previews of each view in a thumbnail at the top of the current view, allowing users to quickly glance at a view or decide whether it is worth switching to another view.

The *ARgus* interface provides three buttons for transitioning between views. Hovering the mouse over a button displays the preview thumbnail of the corresponding view. Clicking on the button activates the view. We used a trajectory and field-of-view interpolation of the camera in the virtual scene when switching views to avoid abrupt transition and disorientation. 3D navigation in the virtual scene is possible by using the mouse. The virtual scene also offers an alternative to preview and switch between views by hovering over and clicking on dedicated 3D widgets: the AR user head for the Headset View and the 3D model of the external camera for the External View. When viewing one of the two augmented video views, the remote user can still use the mouse to navigate in 3D, but this immediately switches the representation to the Virtual View. We also provided a pointing stick and annotation features to enhance communication. The pointing stick can be activated in any view, but it temporarily freezes the Headset View to allow for accurate pointing. Annotations are represented by colored spheres visible in any view.

We conducted a second user study that observed how 12 participants used *ARgus* to provide remote instructions for an AR furniture arrangement task. We

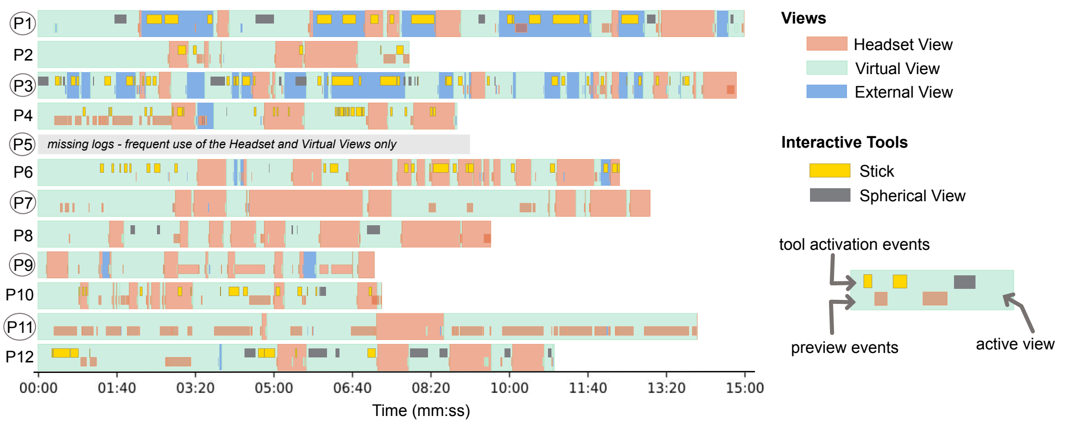


Figure 4.19: Timelines showing the use of the three views, the previews and the pointing stick for each participant when performing the experimental task with *ARgus*. Circled participants were exposed to *ARgus* first.

compared *ARgus* to a controlled condition using only the HEADSET VIEW without any interaction functionalities. We designed a furniture arrangement task, which involves both physical and virtual objects in the 3D scene. Participants had to give instructions to an experimenter, who acted as a confederate, for positioning physical miniature furniture in a virtual 3D model of a house. Participants were given a set of constraints to fulfill for the furniture arrangement. These constraints rely on random parameters to create various arrangement tasks unknown to the experimenter. Participants performed a distinct task for each video condition, and then answered a questionnaire to provide feedback about the two conditions.

To analyze the strategies used by participants to complete the task with *ARgus*, we created timelines showing the time spent in each view, as well as the use of previews and communication tools (Figure 4.19). Since we did not encourage participants to be fast, the time range does not always reflect active collaboration, as some participants spent initial time thinking about the task or exploring the 3D house model. Overall, participants frequently transitioned between views or used previews, which demonstrates that *ARgus* is especially useful to perform the task. This is confirmed by the fact that *ARgus* reduced participants' reliance on verbal instructions and was generally preferred compared to the condition using only the HEADSET VIEW. Nevertheless, we observed that participants employed different strategies when using *ARgus*. While 3 participants (P1, P3, P9) found the EXTERNAL VIEW very useful, others judged that the VIRTUAL VIEW and HEADSET VIEW were enough to complete the task. A few participants, such as P11, relied extensively on the preview feature, whereas others used it temporarily, mainly before switching views. We hypothesized that mastering all combinations of views and previews, as well as developing strategies to use them effectively in various collaboration steps, may require a long learning process that was not assessed in this study.

In summary, we explored how different views can enable a remote desktop user to collaborate with an AR user by perceiving both the physical and virtual content surrounding this AR user. We first compared three representations of the AR workspace and showed that each of them presents different benefits, targeting different collaboration aspects. Based on these findings, we developed *ARgus*, a multi-view collaboration system that provides tools for effectively switching between views and navigating in the AR workspace. A second user study suggests that the flexibility of *ARgus* allows remote users to verify spatial constraints more efficiently and reduces their reliance on verbal instructions. Future work needs to examine the multi-view collaboration strategies from the perspective of the AR user, and how to provide awareness about the visual perception and interaction capabilities of the remote user. Moreover, *ARgus* could be extended to multiple remote users, but this will pose significant challenges in terms of awareness since users will now have distinct representations of the 3D scene and various viewpoints.

4.3 CONCLUSION

Remote collaboration across large interactive spaces is becoming crucial in many situations, allowing remote experts to combine their expertise and offering the flexibility to work from home or reduce travel. As a consequence, new collaborative systems need to handle a wide range of collaboration scenarios and support users with asymmetric device configurations. In this chapter, I first studied the technical aspects of connecting remote users across heterogeneous interactive platforms. I then explored various telepresence systems for enhancing awareness among such remote users with video-mediated communication.

In the first section, I presented several systems for synchronizing CAD data across remote locations, transmitting spatialized 3D audio and reconstructing live 3D head models of remote users. However, these systems are still preliminary proofs of concept at this stage, and would benefit from further development and evaluation on real-life collaboration scenarios. For example, our architecture for synchronizing CAD data could be tested in a large-scale collaboration setting involving multiple design team members interacting with a wide range of immersive and non-immersive devices. The spatialized audio system could be integrated and tested with the telepresence systems proposed in the second section of this chapter.

In the second section, I investigated the potential of video-mediated communication to enhance remote collaboration in different configurations. I first focused on a one-to-one collaboration between two remote collaborators using similar interactive platforms. Next, I explored what happens when a user is away from the work team, involving a one-to-many collaboration. Finally, I addressed the situation in which users do not have access to the same equipment, leading to a one-to-one collaboration between users with immersive and non-immersive devices. For each collaboration configuration, our work is grounded in experimental findings that provide fundamental insights on how users can collaborate through video. In particular, we assessed users' ability to interpret deictic gestures through video and the impact of different representations of augmented reality content on collaboration. Based on these findings, we designed several telepresence systems following the requirements of each collaboration scenarios. I want to emphasize that these systems also represent technical achievements in themselves, involving complex features such as streaming multiple video feeds, augmenting video with virtual content, or transmitting video along with users' positions and actions. As the final step, we evaluated these systems on various collaborative tasks. In multiple cases, we observed that effectively conveying appropriate non-verbal cues or providing useful communication tools can enhance collaboration and reduce the reliance on verbal communication. Nevertheless, conducting an exhaustive evaluation of such collaborative systems is challenging due to the multitude of situations, distinct user roles, and diverse phases in the collaboration process. Therefore, further evaluation of our systems in different tasks and contexts would be welcome.

While most of the work presented in this chapter can be extended to multiple users and multiple remote locations, they have only been tested with two users in one-to-one collaboration. This is mainly due to the complexity of conducting controlled experiments with more than two users, as a larger number of users considerably increases the potential biases of the experiments. However, we need to find new ways to assess collaboration in such multi-user scenarios. Relying on observational studies, as we did to evaluate *ARgus*, may be a solution. It could also be useful to compute real-time indicators of collaboration quality by automatically analyzing users' speech, gaze direction and relative movements, instead of doing it manually after the experiments, as we did in most of our work. In future work, I want to assess how the proposed systems can handle multiple users in each large interactive space. A first step will be to assess *CamRay* with two users in front of each wall-sized display. I also plan to extend these systems to multiple remote locations. For example, *GazeLens* and *ARgus* could be easily extended to integrate multiple remote users. A long-term objective will be to target true hybrid collaboration situations involving both co-located and remote users interacting with heterogeneous devices.

In its current state, this research treats collaboration as a single, simple activity between users. However, collaboration is much more complex, involving different collaboration styles that evolve over time. These styles can include tightly coupled and loose collaboration, subgroup collaboration, as well as spontaneous or side discussions. In future work, I want to assess how the proposed systems can support these different collaboration styles. I also want to extend these systems to better allow transitions among the different phases of a collaboration. For example, *CamRay* can probably handle different collaboration phases with its two video modes, but a solution to transition between the two modes would be required to support various collaboration dynamics.

5

FUTURE PERSPECTIVES AND CLOSING REMARKS

The previous chapters have presented my past work, which focused on investigating individual and collaborative interaction in shared interactive spaces (Chapter 3) and connecting remote users across large interactive spaces through appropriate communication and awareness cues (Chapter 4). This final chapter describes my future research directions and concludes this manuscript with more general remarks.

5.1 INTERACTION ALL ALONG THE MIXED REALITY CONTINUUM

Chapter 3 presented interaction and collaboration techniques targeting different levels of the mixed reality continuum, defined by Milgram et al. [Mil+95] and recently revisited by Skarbez et al. [SSW21]. These techniques include touch interaction on a 2D wall-sized display, 3D gestures in an augmented reality space and haptic interaction in an immersive virtual reality system. However, these interaction techniques remain designed for a specific device and cannot be applied at other levels of the mixed reality continuum. In this section, the term "level" designates a specific position of the continuum, without any notion of discrete tiers or hierarchy.

In general, with the wide diversification of computing devices, solutions exist for visualizing content and interacting at each level of the mixed reality continuum. Additionally, new devices start to offer the ability to transition along this continuum. For example, video see-through headsets now enable users to switch from real vision to various AR views and fully immersive VR views. Despite this, applications and interaction techniques stay siloed at specific levels of the continuum.

As each level has its own benefits and they all complement each other, I argue that new interactive systems should provide users with the ability to interact at multiple levels of the continuum and transition among them. In particular, I believe that such transitions are mandatory to integrate mixed reality into the everyday work pipeline. To be usable, such interactive systems must provide users with consistent interaction techniques to avoid confusing them by changing techniques every time they change level. My future work will concentrate on two main challenges for this research axis: (i) designing large interactive spaces that support transitions along the mixed reality continuum, and (ii) providing consistent interaction techniques that enable users to seamlessly interact across multiple levels.

Supporting transitions along the mixed reality continuum. To achieve these transitions, I envision that users could use either multiple devices or a single device allowing such transitions. I illustrate this concept by presenting a realistic scenario that highlights the benefits of each level of the continuum. This scenario was inspired by observations of real designers in the automotive industry who use a desktop computer along with a VR headset: they use CAD software on the desktop computer to design products and review them in 3D using the VR headset.

This scenario involves engineers who need to analyze a large number of 3D numerical simulation results (Figure 5.1). For tasks such as parameterizing the sim-

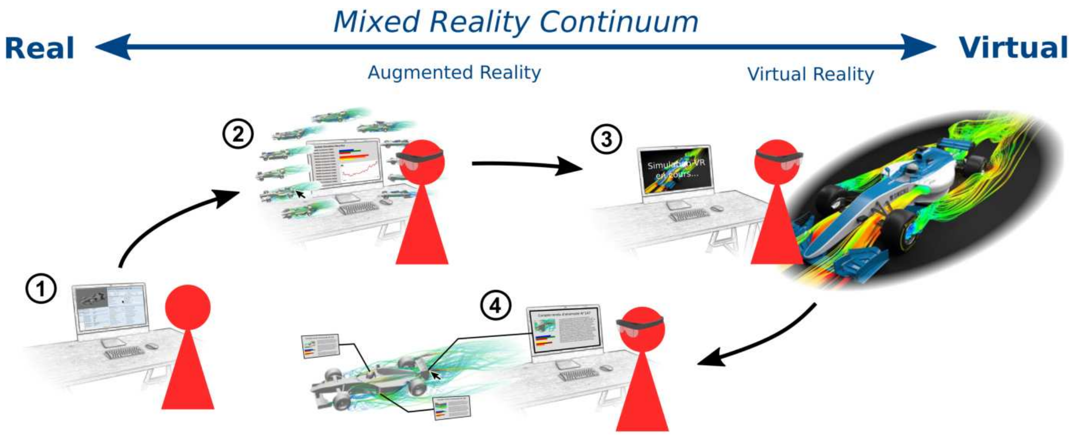


Figure 5.1: Scenario involving interaction along the mixed reality continuum: (1) interaction with the 2D desktop interface, (2) augmentation of the computer screen with 3D models, (3) immersion in a 3D model with virtual reality and (4) interaction with a hybrid interface combining augmented reality and desktop interface.

ulation software or sorting the results, a standard computer allows them to perform these actions efficiently by using the mouse and keyboard (Step 1 in the figure). To compare several simulation results, they use a mixed reality headset that displays these results in 3D around their computer screen (Step 2). In this configuration, they can interact with the data using the mouse both on the screen and in the 3D space, as proposed by Plasson et al. [PBN22]. Later, to better understand an unexpected detail in one of the results, they choose to immerse themselves in a 3D virtual environment, enabling them to view the simulation in context (Step 3). They can thus physically move around and interact in 3D with the displayed data. Finally, once they identify the problem, they return to a hybrid view that combines the computer screen with a 3D view of this specific simulation result. This configuration makes it easy to annotate the data with the keyboard (Step 4).

This scenario illustrates a global vision involving both multi-device interactions and transitions between different levels of the mixed reality continuum. Roo and Hachet [RH17] proposed *One Reality*, a conceptual framework enabling a comparable scenario. *One Reality* allows users to interact with a physical object and its virtual counterpart at 6 levels of the continuum, ranging from the physical object alone to an immersive view of the virtual counterpart in VR. Although all levels are synchronized, users still need to switch devices to transition between certain levels. Such systems require a distributed software architecture to share data across multiple devices. This architecture could be inspired by the work achieved during my PhD to distribute data across VR devices in a collaborative context [**Fle+10**b].

As a first step, I plan to explore simpler configurations that combine desktop interfaces, mobile devices or wall-sized displays with mixed reality technologies. For example, in the context of co-located collaboration, James et al. [JBC23] propose to extend a wall-sized display with shared and personal surfaces displayed using AR headsets. I have also initiated a project on 3D editing, in which we want to augment a standard computer screen by incorporating 3D views positioned around the screen with augmented reality. The goal of this future work will be to gain insights on how users can interact at various levels of the continuum and assess the need for transitions among these levels.

Providing consistent interaction across the mixed reality continuum. To seamlessly interact across different levels of the mixed reality continuum, users need consistent techniques that prevent them from having to switch to a whole new set of interaction techniques every time they change level. A first solution is to enable users to interact at multiple levels using the same input device and interaction technique. For example, Plasson et al. [PBN22] propose to use a mouse for interacting on a 2D screen, as well as with 3D views displayed next to the screen through an AR headset. James et al. [JBC23] extend the pointing and grabbing techniques from an AR headset to also grab 2D content on a wall-sized display. I plan to further explore multi-level interaction techniques for manipulating both 2D and 3D content with various devices and different interactive setups.

However, we cannot expect users to interact with the same technique across all levels of the continuum, given the wide range of devices available. We must therefore enable them to change techniques without being lost or having to relearn all the interaction mechanisms for each new application. I think that efforts should be made to propose standardized interaction techniques, especially when it comes to 3D interaction, which is fairly new to users and remains very specific from one application to another. Interaction discoverability should also be improved in mixed reality systems, as it is the case for 2D interfaces. Finally, we have to keep in mind that the new techniques must allow multiple users to interact in the same interactive space and potentially support collaborative activities.

5.2 HYBRID COLLABORATION ACROSS LARGE INTERACTIVE SPACES

Chapter 3 explored various co-located collaboration scenarios, while Chapter 4 focused on remote collaboration. However, none of this work investigates real hybrid collaborative situations, including both co-located and remote users. While *GazeLens* (Section 4.2.2) aimed to integrate a remote user into a co-located collaboration, this work did not explore the co-located aspect of the collaboration.

Hybrid collaboration has become a necessity due to major changes in our society and the new organization of work. We are experiencing more and more diverse collaborative situations, such as meetings with a colleague working from home, or work sessions between two distant groups. The COVID-19 pandemic significantly accentuated this trend [Yan+22]. However, current computer-mediated collaboration systems often lack flexibility to adequately support hybrid collaboration. This can lead to awkward situations in which colleagues within the same building opt to stay in their individual office for a videoconference meeting instead of attending together, or are forced to have side conversations via chat during such meetings.

My future research aims to investigate how large interactive spaces can foster real-time collaboration in hybrid situations. For example, these situations may include remote collaboration among co-located subgroups or collaboration between a co-located group and multiple remote users (Figure 5.2). As suggested in the previous research axis, users may interact at different levels of the mixed reality continuum using heterogeneous devices, ranging from simple smartphones to immersive VR rooms. I think that this diversity of devices has the potential to enhance hybrid collaboration. Nevertheless, this raises new challenges regarding (i) how to integrate users with heterogeneous devices in the collaboration, (ii) how to provide appropriate awareness among users, regardless of whether they

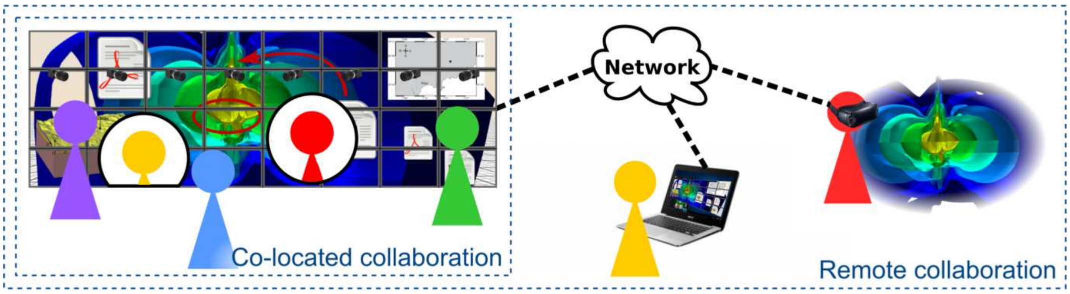


Figure 5.2: Hybrid collaboration involving co-located and remote users interacting with heterogeneous devices.

are co-located or remote, and (iii) how to support different dynamics during the collaboration process. Without trying to mimic collaboration in the real world, I believe we need new ways of collaborating that take advantage of the specific capabilities of each interaction device, in a similar spirit as the "Beyond being there" concept described by Hollan & Stornetta [HS92].

Integrating users with heterogeneous devices. As presented in the previous research axis, I envision that users will be able to interact at multiple levels of the mixed reality continuum and transition among these levels. Consequently, users will collaborate with both co-located and remote collaborators across various levels of the continuum. These heterogeneous situations provide many opportunities for exploring new collaboration scenarios. For instance, given that using immersive VR for extended periods of time can be strenuous, users could use non-immersive devices to remotely monitor collaborators immersed in VR and interact with them. This approach allows users to reserve VR for specific tasks during long work sessions, taking turns in VR as needed. I also plan to study another scenario in which a co-located group collaborates in front of a wall-sized display, while remote collaborators, who do not have access to such equipment, join them using VR headsets that display a virtual version of the content. A few studies have explored asymmetric collaboration across heterogeneous devices, but they mainly focused on co-located collaboration. For example, *ShareVR* [Gug+17] and *TransceiVR* [Tho+20] use a smartphone or a tablet to interact with a user immersed in VR, while *ShARe* [Jan+20] and *HMD Light* [Wan+20b] use a projector mounted on an AR or VR headset to share the view of the headset user. I plan to extend this previous work to broader hybrid collaborative situations including remote users.

The main challenge is to provide all users with appropriate interaction techniques to act on shared content and communicate their ideas, regardless of their geographical location or their device. These techniques should leverage the potential of every device to allow users to have complementary interaction capabilities. Additionally, it is necessary to find appropriate ways to represent users' activities and interaction capabilities to improve understanding among them. Given that users may have varying interaction capabilities, it is crucial that they can understand what others are currently doing and what they can do to enable effective collaboration. As a first step, I plan to extend *ARgus* (Section 4.2.3) to support multiple co-located AR users, as well as multiple remote users. This will require finding solutions to allow AR users to accurately understand what is the viewpoint of each remote user on the AR workspace, and give all remote users the ability to participate in 3D interaction.

Providing appropriate awareness between co-located and remote users. Allowing effective collaboration in hybrid situations requires to provide appropriate awareness among all users. This awareness is crucial for enhancing their mutual understanding and helping them build a common ground [CB91]. Providing such awareness is challenging between co-located and remote users, as they do not share the same interactive space and cannot see each other directly. I plan to explore multiple solutions for representing the remote users, but also the space surrounding them, in a way that ensures seamless interaction between co-located and remote users. Representing the space surrounding users is especially important to facilitate the establishment of a common ground, as we studied in *ARgus* (Section 4.2.3).

I first want to explore different visual representations of the remote users in collaborative situations involving mixed reality technologies. There is not clear consensus regarding the impact of user representations on collaboration. Yoon et al. [Yoo+19] compared the effects of realistic and cartoon-like avatars on social presence, while Congdon et al. [Con+23] compared the effects of video and 3D avatar representations on trust. Both studies concluded that the results could highly depend on the collaborative context and the environment surrounding the users. Although solutions exist to create high-quality realistic avatars, such as *Pixel Codec Avatars* [Ma+21], I believe that there are some collaborative situations where avatars may not be the most appropriate representation. It is especially the case for collaborative situations including users with both immersive and non-immersive devices, as using avatars for non-immersive users may not be meaningful. In such situations, I want to experiment with solutions that integrate real video streams into virtual environments in a way that goes beyond real-world collaboration. For example, we could attach virtual windows displaying remote collaborators' video on the side of the users' field of view in a VR environment or on the users' wrist in an AR environment. This will create a virtual face-to-face with the remote collaborators, as we explored with *CamRay* on wall-sized displays (Section 4.2.1). We also need to find solutions to represent users of immersive devices for the collaborators using non-immersive devices.

Although showing the spaces surrounding remote users is straightforward in non-immersive contexts with cameras, it becomes challenging for mixed reality environments that overlap multiple remote spaces with both physical and virtual content. Most previous work on remote collaboration in mixed reality focused on host-guest situations, where the guest is immersed in the augmented environment of the host [Teo+19; Piu+19; Bai+20]. Other research proposed sharing only a few physical objects [Ort+16] or virtual content [Mah+19], but not the entire spaces surrounding users. However, some collaborative situations require a shared space that combines the spaces of all remote users with their corresponding physical constraints. A few studies have explored this aspect, mainly focusing on the technical aspects of reconstructing and blending users' physical spaces [LMR14]. I plan to approach the problem from a different angle by studying how users perceive the remote spaces of their collaborators and identifying which cues are mandatory to build a mental representation of the shared space. I will then investigate representations that mix symbolic and realistic elements to reveal this shared space. These representations should prevent users from perceiving the shared space as a superposition of individual spaces. Instead, they should facilitate the establishment of a common ground between users, enhancing their mutual understanding.

Supporting various collaboration dynamics. As the number of users involved in hybrid collaboration increases, not all users will be collaborating together at all times. Collaborative systems will thus have to support different moments in collaboration, such as tightly coupled and loose collaboration, subgroup collaboration, and spontaneous or side discussions. However, current telepresence systems, including those presented in Section 4.2, do not adequately support these dynamic collaboration scenarios.

As an initial step, I plan to study collaboration dynamics in co-located situations without technology mediation. In particular, I want to observe groups of co-located users interacting in large interactive spaces, such as rooms equipped with large screens, tabletops, or AR devices. Building upon these observations, the goal is to extend our work on telepresence systems for wall-sized displays (Section 4.2.1) to support these collaboration dynamics. An obvious first goal is to explore manual and automatic solutions for switching between the *Follow-Local* and *Follow-Remote* modes of *CamRay*. Moreover, I can imagine various other collaborative feature based on video. For example, some video windows could appear or disappear as appropriate to manage tightly coupled versus loose collaboration. Other video windows could be split and displayed at different positions to encourage users to move to specific areas of the screens, thus fostering subgroup collaboration across remote platforms. Additional devices, such as smartphones, could also be used on the fly to make side conversations possible. In addition to video, I believe it would be valuable to incorporate spatialized 3D audio by using the system we developed (Section 4.1.2). Spatialized 3D audio could enable users to determine remote collaborator positions, manage subgroup discussions without disturbing others or ensure privacy for side conversations.

5.3 CLOSING REMARKS

The demand for computer-supported cooperative work has never been more critical, given the substantial growth of digital data and significant societal changes, including new work organization and the green transition. An increasing number of individuals are required to work from home or collaborate with colleagues worldwide while limiting their travel to mitigate their environmental footprint. Although computer-supported cooperative work has been studied for several decades, the vast majority of previous work considered simplistic collaborative situations, such as one-to-one or group collaboration among individuals with similar roles. However, real-world collaboration is considerably more complex, involving multiple roles, users entering and leaving the collaboration, and different collaboration dynamics, such as spontaneous or side discussions. I argue in this manuscript that large interactive spaces provide a unique opportunity to support such complex collaborative situations across time and space. Nevertheless, further research is still needed to handle hybrid collaboration and transitions along the mixed reality continuum.

While my contributions and future perspectives mainly concentrate on synchronous collaboration, large interactive spaces hold significant potential for fostering collaboration in asynchronous situations. This is a typical case where computer systems can provide users with collaborative interaction that goes far beyond what is possible without technology mediation. Fender and Holz [FH22] illustrate the benefits of mixed reality technology for co-located asynchronous collaboration.

However, only a few studies have addressed asynchronous collaboration, as observed by Irlitti et al. [Irl+16], leaving plenty of space for design exploration, as suggested by Chow et al. [Cho+19]. I believe that asynchronous collaboration can be a promising long-term perspective for future work.

Targeting complex collaborative situations also raises the question of how to evaluate collaboration, as it cannot be solely assessed through performance measures in lab experiments. The success of collaboration is determined by many underlying indicators that are challenging to quantify. These indicators include social presence, mutual understanding, active participation, and feeling of closeness or friendliness among collaborators. A few studies have attempted to measure some of these indicators through questionnaires or post-experiment conversational analysis, as we have done in some of our work [**Ave+17**; **Oku+20**; **FFT22b**]. However, these analyses are difficult and time-consuming, thus limiting the number of indicators that can be measured. With current advances in sensing technologies and artificial intelligence, we just started to create a system that evaluates collaboration quality in real time [**Léc+23**], as part of the PhD work of A. Léchappé, co-supervised with M. Cholet and C. Dumas, and in collaboration with A. Milliat. At the current stage, this system can collect gaze and speech signals, as well as compute speaking time distribution, turn-taking, speech overlaps, joint visual attention, and mutual gaze. A first iteration used these indicators to differentiate situations with active collaboration from those without collaboration. Future steps will consist of detecting more complex collaborative situations, and providing users with real-time feedback to prevent critical situations arising from poor collaboration.

To conclude, I believe that large interactive spaces hold huge potential for fostering collaboration in various real-world situations. Nevertheless, many challenges persist in providing collaborators with rich social interaction and appropriate collaborative features. Close collaboration with researchers in social sciences will be crucial to better understand how individuals collaborate through technology and to adequately evaluate such collaboration.

BIBLIOGRAPHY

[AAT13] Matt Adcock, Stuart Anderson, and Bruce Thomas. “RemoteFusion: Real Time Depth Camera Fusion for Remote Collaboration on Physical Tasks.” In: *Proceedings of the SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*. VRCAI ’13. Hong Kong, Hong Kong: ACM, 2013, pp. 235–242. DOI: `10.1145/2534329.2534331`.

[ADL09] Laurent Aguerreche, Thierry Duval, and Anatole Lécuyer. “3-Hand Manipulation of Virtual Objects.” In: *Proceedings of the Joint Virtual Reality Conference of EGVE - ICAT - EuroVR*. The Eurographics Association, 2009. DOI: `10.2312/EGVE/JVRC09/153-156`.

[ADL10a] Laurent Aguerreche, Thierry Duval, and Anatole Lécuyer. “Comparison of Three Interactive Techniques for Collaborative Manipulation of Objects in Virtual Reality.” In: *Computer Graphics International*. CGI’ 10. 2010. URL: `https://inria.hal.science/inria-00534087`.

[ADL10b] Laurent Aguerreche, Thierry Duval, and Anatole Lécuyer. “Reconfigurable Tangible Devices for 3D Virtual Object Manipulation by Single or Multiple Users.” In: *Proceedings of the Symposium on Virtual Reality Software and Technology*. VRST ’10. Hong Kong: ACM, 2010, pp. 227–230. DOI: `10.1145/1889863.1889913`.

[Ahs+21] Tooba Ahsen, Zi Yi Lim, Aaron L. Gardony, Holly A. Taylor, Jan P de Ruiter, and Fahad Dogar. “The Effects of Network Outages on User Experience in Augmented Reality Based Remote Collaboration - An Empirical Study.” In: *Proceedings of the ACM on Human-Computer Interaction* 5.CSCW2 (2021). DOI: `10.1145/3476054`.

[Akk+16] Deepak Akkil, Jobin Mathew James, Poika Isokoski, and Jari Kangas. “GazeTorch: Enabling Gaze Awareness in Collaborative Physical Tasks.” In: *Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems*. CHI EA ’16. San Jose, California, USA: ACM, 2016, pp. 1151–1158. DOI: `10.1145/2851581.2892459`.

[ATI18] Deepak Akkil, Biju Thankachan, and Poika Isokoski. “I See What You See: Gaze Awareness in Mobile Video Collaboration.” In: *Proceedings of the Symposium on Eye Tracking Research & Applications*. ETRA ’18. Warsaw, Poland: ACM, 2018. DOI: `10.1145/3204493.3204542`.

[AEN10] Christopher Andrews, Alex Endert, and Chris North. “Space to Think: Large High-Resolution Displays for Sensemaking.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’10. Atlanta, Georgia, USA: ACM, 2010, pp. 55–64. DOI: `10.1145/1753326.1753336`.

[And+11] Christopher Andrews, Alex Endert, Beth Yost, and Chris North. "Information Visualization on Large, High-Resolution Displays: Issues, Challenges, and Opportunities." In: *Information Visualization* 10.4 (2011), pp. 341–355. DOI: 10.1177/1473871611415997.

[AG00] Rajarathinam Arangarasan and Rajit Gadh. "Geometric Modeling and Collaborative Design in a Multi-Modal Multi-Sensory Virtual Environment." In: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference Volume 1: 20th Computers and Information in Engineering Conference (2000), pp. 757–765. DOI: 10.1115/DETC2000/CIE-14592.

[Ara+13] Bruno R. De Araújo, Géry Casiez, Joaquim A. Jorge, and Martin Hachet. "Mockup Builder: 3D modeling on and above the surface." In: *Computers & Graphics* 37.3 (2013), pp. 165–178. DOI: 10.1016/j.cag.2012.12.005.

[AA13] Ferran Argelaguet and Carlos Andujar. "A survey of 3D object selection techniques for virtual environments." In: *Computers & Graphics* 37.3 (2013), pp. 121–136. DOI: 10.1016/j.cag.2012.12.003.

[AFB15] Ignacio Avellino, Cédric Fleury, and Michel Beaudouin-Lafon. "Accuracy of Deictic Gestures to Support Telepresence on Wall-Sized Displays." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI'15. Seoul, Republic of Korea: ACM, 2015, pp. 2393–2396. DOI: 10.1145/2702123.2702448.

[Ave+17] Ignacio Avellino, Cédric Fleury, Wendy E. Mackay, and Michel Beaudouin-Lafon. "CamRay: Camera Arrays Support Remote Collaboration on Wall-Sized Displays." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI'17. Denver, Colorado, USA: ACM, 2017, pp. 6718–6729. DOI: 10.1145/3025453.3025604.

[Bai+20] Huidong Bai, Prasanth Sasikumar, Jing Yang, and Mark Billinghurst. "A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '20. Honolulu, HI, USA: ACM, 2020, pp. 1–13. DOI: 10.1145/3313831.3376550.

[Bal+99] Ravin Balakrishnan, George Fitzmaurice, Gordon Kurtenbach, and William Buxton. "Digital Tape Drawing." In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST '99. Asheville, North Carolina, USA: ACM, 1999, pp. 161–169. DOI: 10.1145/320719.322598.

[BNB07] Robert Ball, Chris North, and Doug A. Bowman. "Move to Improve: Promoting Physical Navigation to Increase User Performance with Large Displays." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '07. San Jose, California, USA: ACM, 2007, pp. 191–200. DOI: 10.1145/1240624.1240656.

[Bar+21] Andrea Bartl, Stephan Wenninger, Erik Wolf, Mario Botsch, and Marc Erich Latoschik. "Affordable But Not Cheap: A Case Study of the Effects of Two 3D-Reconstruction Methods of Virtual Humans." In: *Frontiers in Virtual Reality* 2 (2021). DOI: 10.3389/frvir.2021.694617.

[Bea11] Michel Beaudouin-Lafon. "Lessons Learned from the WILD Room, a Multisurface Interactive Environment." In: *Proceedings of the Conference on l'Interaction Homme-Machine*. IHM '11. Sophia Antipolis, France: ACM, 2011. DOI: 10.1145/2044354.2044376.

[BM00] Michel Beaudouin-Lafon and Wendy E. Mackay. "Reification, Polymorphism and Reuse: Three Principles for Designing Visual Interfaces." In: *Proceedings of the Working Conference on Advanced Visual Interfaces*. AVI '00. Palermo, Italy: ACM, 2000, pp. 102–109. DOI: 10.1145/345513.345267.

[Bea+12] Michel Beaudouin-Lafon et al. "Multisurface Interaction in the WILD Room." In: *Computer* 45.4 (2012), pp. 48–56. DOI: 10.1109/MC.2012.110.

[Bec+13] Stephan Beck, André Kunert, Alexander Kulik, and Bernd Froehlich. "Immersive Group-to-Group Telepresence." In: *IEEE Transactions on Visualization and Computer Graphics* 19.4 (2013), pp. 616–625. DOI: 10.1109/TVCG.2013.33.

[Bee+10] Thabo Beeler, Bernd Bickel, Paul Beardsley, Bob Sumner, and Markus Gross. "High-Quality Single-Shot Capture of Facial Geometry." In: *ACM ACM Transactions on Graphics (siggraph'10)* 29.3 (2010), 40:1–9. DOI: 10.1145/1778765.1778777.

[Bee+11] Thabo Beeler, Fabian Hahn, Derek Bradley, Bernd Bickel, Paul Beardsley, Craig Gotsman, Robert W. Sumner, and Markus Gross. "High-quality passive facial performance capture using anchor frames." In: *ACM Transactions on Graphics (siggraph'11)* 30.4 (2011), 75:1–10. DOI: 10.1145/2010324.1964970.

[Ben+95] Steve Benford, John Bowers, Lennart E. Fahlén, Chris Greenhalgh, and Dave Snowdon. "User Embodiment in Collaborative Virtual Environments." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '95. Denver, Colorado, USA: ACM Press/Addison-Wesley Publishing Co., 1995, pp. 242–249. DOI: 10.1145/223904.223935.

[Ber99] Julien Berta. "Integrating VR and CAD." In: *IEEE Computer Graphics and Applications* 19.5 (1999), pp. 14–19. DOI: 10.1109/38.788793.

[BB05a] Anastasia Bezerianos and Ravin Balakrishnan. "The Vacuum: Facilitating the Manipulation of Distant Objects." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '05. Portland, Oregon, USA: ACM, 2005, pp. 361–370. DOI: 10.1145/1054972.1055023.

[BB05b] Anastasia Bezerianos and Ravin Balakrishnan. "View and space management on large displays." In: *IEEE Computer Graphics and Applications* 25.4 (2005), pp. 34–43. DOI: 10.1109/MCG.2005.92.

[BB09] Xiaojun Bi and Ravin Balakrishnan. “Comparing Usage of a Large High-Resolution Display to Single or Dual Desktop Displays for Daily Work.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’09. Boston, MA, USA: ACM, 2009, pp. 1005–1014. DOI: 10.1145/1518701.1518855.

[BCL15] Mark Billinghurst, Adrian Clark, and Gun Lee. “A Survey of Augmented Reality.” In: *Foundations and Trends® in Human–Computer Interaction* 8.2-3 (2015), pp. 73–272. DOI: 10.1561/1100000049.

[BK02] Mark Billinghurst and Hirokazu Kato. “Collaborative Augmented Reality.” In: *Communication of the ACM* 45.7 (2002), pp. 64–70. DOI: 10.1145/514236.514265.

[Bil+02] Mark Billinghurst, Hirokazu Kato, Kiyoshi Kiyokawa, Daniel Belcher, and Ivan Poupyrev. “Experiments with Face-To-Face Collaborative AR Interfaces.” In: *Virtual Reality* 6.3 (2002), pp. 107–121. DOI: 10.1007/s100550200012.

[BWF98] Mark Billinghurst, Suzanne Weghorst, and Thomas Furness. “Shared Space: An Augmented Reality Approach for Computer Supported Collaborative Work.” In: *Virtual Reality* 3.1 (1998), pp. 25–36. DOI: 10.1007/BF01409795.

[Blo22] Nick Bloom. *Working from Home and the Future of U.S. Economic Growth under COVID*. https://youtu.be/jtdFIZx3hyk, accessed: 2022-03-23. 2022.

[BHI93] Sara A. Bly, Steve R. Harrison, and Susan Irwin. “Media Spaces: Bringing People Together in a Video, Audio, and Computing Environment.” In: *Commun. ACM* 36.1 (1993), pp. 28–46. DOI: 10.1145/151233.151235.

[Bou+10] Patrick Bourdot, Thomas Convard, Flavien Picon, Mehdi Ammi, Damien Touraine, and Jean-Marc Vézien. “VR–CAD integration: Multimodal immersive interaction and advanced haptic paradigms for implicit edition of CAD models.” In: *Computer-Aided Design* 42.5 (2010). Advanced and Emerging Virtual and Augmented Reality Technologies in Product Design, pp. 445–461. DOI: 10.1016/j.cad.2008.10.014.

[BT02] Patrick Bourdot and Damien Touraine. “Polyvalent display framework to control virtual navigations by 6DOF tracking.” In: *Proceedings of the IEEE Virtual Reality (VR)*. 2002, pp. 277–278. DOI: 10.1109/VR.2002.996537.

[Bux09] Bill Buxton. “Mediaspace – Meaningspace – Meetingspace.” In: *Media Space 20 + Years of Mediated Life*. Ed. by Steve Harrison. London: Springer, 2009, pp. 217–231. DOI: 10.1007/978-1-84882-483-6_13.

[Bux+00] William Buxton, George Fitzmaurice, Ravin Balakrishnan, and Gordon Kurtenbach. “Large displays in automotive design.” In: *IEEE Computer Graphics and Applications* 20.4 (2000), pp. 68–75. DOI: 10.1109/38.851753.

[Bux92] William A. S. Buxton. “Telepresence: Integrating Shared Task and Person Spaces.” In: *Proceedings of the Conference on Graphics Interface, GI’92*. Vancouver, British Columbia, Canada: Morgan Kaufmann Publishers Inc., 1992, pp. 123–129. DOI: 10.5555/155294.155309.

[Cai+10] Qin Cai, David Gallup, Cha Zhang, and Zhengyou Zhang. “3D Deformable Face Tracking with a Commodity Depth Camera.” In: *Proceedings of the European Conference on Computer Vision Conference on Computer Vision: Part III*. ECCV’10. Heraklion, Crete, Greece: Springer-Verlag, 2010, pp. 229–242. DOI: 10.1007/978-3-642-15558-1_17.

[CM92] Thomas Caudell and David Mizell. “Augmented reality: an application of heads-up display technology to manual manufacturing processes.” In: *Proceedings of the Hawaii International Conference on System Sciences*. Vol. ii. 1992, 659–669 vol.2. DOI: 10.1109/HICSS.1992.183317.

[CBF14] Olivier Chapuis, Anastasia Bezerianos, and Stelios Frantzeskakis. “Smarties: An Input System for Wall Display Development.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’14. Toronto, Ontario, Canada: ACM, 2014, pp. 2763–2772. DOI: 10.1145/2556288.2556956.

[CF17] Haiwei Chen and Henry Fuchs. “Supporting Free Walking in a Large Virtual Environment: Imperceptible Redirected Walking with an Immersive Distractor.” In: *Proceedings of the Computer Graphics International Conference*. CGI ’17. Yokohama, Japan: ACM, 2017. DOI: 10.1145/3095140.3095162.

[Che02] Milton Chen. “Leveraging the Asymmetric Sensitivity of Eye Contact for Videoconference.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’02. Minneapolis, Minnesota, USA: ACM, 2002, pp. 49–56. DOI: 10.1145/503376.503386.

[Che+13] Weiya Chen, Anthony Plancoulaine, Nicolas Férey, Damien Touraine, Julien Nelson, and Patrick Bourdot. “6DoF Navigation in Virtual Worlds: Comparison of Joystick-Based and Head-Controlled Paradigms.” In: *Proceedings of the Symposium on Virtual Reality Software and Technology*. VRST ’13. Singapore: ACM, 2013, pp. 111–114. DOI: 10.1145/2503713.2503754.

[Che+18] Xiang Chen, Ye Tao, Guanyun Wang, Runchang Kang, Tovi Grossman, Stelian Coros, and Scott E. Hudson. “Forte: User-Driven Generative Design.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’18. Montreal QC, Canada: ACM, 2018, pp. 1–12. DOI: 10.1145/3173574.3174070.

[Che+19] Lung-Pan Cheng, Eyal Ofek, Christian Holz, and Andrew D Wilson. “VRoamer: Generating On-The-Fly VR Experiences While Walking inside Large, Unknown Real-World Building Environments.” In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE. 2019, pp. 359–366. DOI: 10.1109/VR.2019.8798074.

[Cho+19] Kevin Chow, Caitlin Coyiuto, Cuong Nguyen, and Dongwook Yoon. “Challenges and Design Considerations for Multimodal Asynchronous Collaboration in VR.” In: *Proceedings of the ACM on Human-Computer Interaction* 3.CSCW (2019). DOI: 10.1145/3359142.

[CB91] Herbert H. Clark and Susan E. Brennan. "Grounding in Communication." In: *Perspectives on Socially Shared Cognition*. Ed. by Lauren Resnick, Levine B., M. John, Stephanie Teasley, and D. American Psychological Association, 1991, pp. 259–292. DOI: 10.1037/10096-006.

[CM81] Herbert H. Clark and Catherine R. Marshall. "Definite Knowledge and Mutual Knowledge." In: *Elements of Discourse Understanding*. Ed. by Aravind K. Joshi, Bonnie L. Webber, and Ivan A. Sag. Cambridge University Press, 1981, pp. 10–63.

[CSB83] Herbert H. Clark, Robert Schreuder, and Samuel Buttrick. "Common ground at the understanding of demonstrative reference." In: *Journal of Verbal Learning and Verbal Behavior* 22.2 (1983), pp. 245–258. DOI: 10.1016/S0022-5371(83)90189-5.

[Cof+13] Dane Coffey, Chi-Lun Lin, Arthur G Erdman, and Daniel F Keefe. "Design by Dragging: An Interface for Creative Forward and Inverse Design with Simulation Ensembles." In: *IEEE Trans. on Visualization and Computer Graphics* 19.12 (2013), pp. 2783–2791. DOI: 10.1109/TVCG.2013.147.

[Cof+12] Dane Coffey, Nicholas Malbraaten, Trung Bao Le, Iman Borazjani, Fotis Sotiropoulos, Arthur G. Erdman, and Daniel F. Keefe. "Interactive Slice WIM: Navigating and Interrogating Volume Data Sets Using a Multisurface, Multitouch VR Interface." In: *IEEE Transactions on Visualization and Computer Graphics* 18.10 (2012), pp. 1614–1626. DOI: 10.1109/TVCG.2011.283.

[CAK93] Michael Cohen, Shigeaki Aoki, and Nobuo Koizumi. "Augmented audio reality: telepresence/VR hybrid acoustic environments." In: *Proceedings of the IEEE International Workshop on Robot and Human Communication*. 1993, pp. 361–364. DOI: 10.1109/ROMAN.1993.367692.

[Con+23] Ben J. Congdon, Gun Woo (Warren) Park, Jingyi Zhang, and Anthony Steed. "Comparing Mixed Reality Agent Representations: Studies in the Lab and in the Wild." In: *Proceedings of the Symposium on Virtual Reality Software and Technology*. VRST '23. Christchurch, New Zealand: ACM, 2023. DOI: 10.1145/3611659.3615719.

[CB04] Thomas Convard and Patrick Bourdot. "History Based Reactive Objects for Immersive CAD." In: *Proceedings of the Symposium on Solid Modeling and Applications*. SM '04. Genoa, Italy: Eurographics Association, 2004, pp. 291–296. DOI: 10.2312/sm.20041404.

[CAC21] Emmanuel Courtoux, Caroline Appert, and Olivier Chapuis. "WallTokens: Surface Tangibles for Vertical Displays." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '21. Yokohama, Japan: ACM, 2021. DOI: 10.1145/3411764.3445404.

[Cru+92] Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon, and John C. Hart. "The CAVE: Audio Visual Experience Automatic Virtual Environment." In: *Communication of the ACM* 35.6 (1992), pp. 64–72. DOI: 10.1145/129888.129892.

[Cze+03] Mary Czerwinski, Greg Smith, Tim Regan, Brian Meyers, George Robertson, and Gary Starkweather. "Toward Characterizing the Productivity Benefits of Very Large Displays." In: *Proceedings of the IFIP TC13 International Conference On Human-Computer Interaction*. INTERACT '03. IOS Press, 2003, pp. 9–16.

[De +13] Bruno R. De Araújo, Géry Casiez, Joaquim A. Jorge, and Martin Hachet. "Mockup Builder: 3D modeling on and above the surface." In: *Computers & Graphics* 37.3 (2013), pp. 165–178. DOI: `10.1016/j.cag.2012.12.005`.

[De 67] Edward De Bono. *The use of lateral thinking*. London, England: Jonathan Cape, 1967.

[Dét06] Françoise Détienne. "Collaborative design: Managing task interdependencies and multiple perspectives." In: *Interacting with Computers* 18.1 (2006), pp. 1–20. DOI: `10.1016/j.intcom.2005.05.001`.

[Dou+12] Mingsong Dou, Ying Shi, Jan-Michael Frahm, Henry Fuchs, Bill Mauchly, and Mod Marathe. "Room-sized informal telepresence system." In: *Proceedings of the IEEE Virtual Reality Workshops (VRW)*. 2012, pp. 15–18. DOI: `10.1109/VR.2012.6180869`.

[DLT06] Thierry Duval, Anatole Lécuyer, and Sébastien Thomas. "SkeweR: a 3D Interaction Technique for 2-User Collaborative Manipulation of Objects in Virtual Environments." In: *Proceedings of the 3D User Interfaces symposium (3DUI'06)*. 2006, pp. 69–72. DOI: `10.1109/VR.2006.119`.

[Elr+92] Scott Elrod et al. "Liveboard: A Large Interactive Display Supporting Group Meetings, Presentations, and Remote Collaboration." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '92. Monterey, California, USA: ACM, 1992, pp. 599–607. DOI: `10.1145/142750.143052`.

[FFT22a] Arthur Fages, Cédric Fleury, and Theophanis Tsandilas. "ARgus : système multi-vues pour collaborer à distance avec un utilisateur en réalité augmentée." In: *Proceedings of the conférence francophone sur l'Interaction Homme-Machine*. Demonstration. Namur, Belgium, 2022. URL: `https://hal.science/hal-03762816`.

[FFT22b] Arthur Fages, Cédric Fleury, and Theophanis Tsandilas. "Understanding Multi-View Collaboration between Augmented Reality and Remote Desktop Users." In: *Proceedings of the ACM on Human-Computer Interaction* 6. CSCW2 (2022). DOI: `10.1145/3555607`.

[FFT23] Arthur Fages, Cédric Fleury, and Theophanis Tsandilas. "Conception collaborative au travers de versions parallèles en réalité augmentée." In: *Proceedings of the Conference on l'Interaction Humain-Machine*. IHM '23. TROYES, France: ACM, 2023. DOI: `10.1145/3583961.3583978`.

[FMS93] Steven Feiner, Blair Macintyre, and Dorée Seligmann. "Knowledge-Based Augmented Reality." In: *Communications of the ACM* 36.7 (1993), pp. 53–62. DOI: `10.1145/159544.159587`.

[FH22] Andreas Rene Fender and Christian Holz. "Causality-Preserving Asynchronous Reality." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '22. New Orleans, LA, USA: ACM, 2022. DOI: `10.1145/3491102.3501836`.

[Fio+02] Michele Fiorentino, Raffaele de Amicis, Giuseppe Monno, and Andre Stork. "Spacedesign: A Mixed Reality Workspace for Aesthetic Industrial Design." In: *Proceedings of the International Symposium on Mixed and Augmented Reality*. ISMAR '02. IEEE Computer Society, 2002, p. 86. DOI: `10.1109/ISMAR.2002.1115077`.

[Fle+04] Timo Fleisch, Gino Brunetti, Pamela Santos, and André Stork. "Stroke-input methods for immersive styling environments." In: *Proceedings of the Shape Modeling Applications conference*. 2004, pp. 275–283. DOI: `10.1109/SMI.2004.1314514`.

[Fle+15a] Cédric Fleury, Ignacio Avellino, Michel Beaudouin-Lafon, and Wendy E. Mackay. "Telepresence Systems for Large Interactive Spaces." In: *Workshop on Everyday Telepresence: Emerging Practices and Future Research Directions at the ACM conference on Human Factors in Computing Systems*. Seoul, South Korea, 2015. URL: `https://hal.science/hal-01242304`.

[Fle+10a] Cédric Fleury, Alain Chauffaut, Thierry Duval, Valérie Gouranton, and Bruno Arnaldi. "A Generic Model for Embedding Users' Physical Workspaces into Multi-Scale Collaborative Virtual Environments." In: *Proceedings of the International Conference on Artificial Reality and Telexistence*. ICAT'10. Adelaide, Australia, 2010. URL: `https://hal.science/inria-00534096`.

[Fle+10b] Cédric Fleury, Thierry Duval, Valérie Gouranton, and Bruno Arnaldi. "A New Adaptive Data Distribution Model for Consistency Maintenance in Collaborative Virtual Environments." In: *Proceedings of the Joint Virtual Reality Conference of EuroVR - EGVE - VEC*. EGVE - JVRC'10. Stuttgart, Germany: Eurographics Association, 2010, pp. 29–36. DOI: `10.2312/EGVE/JVRC10/029-036`.

[Fle+10c] Cédric Fleury, Thierry Duval, Valérie Gouranton, and Bruno Arnaldi. "Architectures and Mechanisms to Maintain efficiently Consistency in Collaborative Virtual Environments." In: *Workshop on Software Engineering and Architectures for Realtime Interactive Systems (SEARIS) at the IEEE Conference on Virtual Reality*. Waltham, United States, 2010. URL: `https://hal.science/inria-00534082`.

[Fle+12] Cédric Fleury, Thierry Duval, Valérie Gouranton, and Anthony Steed. "Evaluation of Remote Collaborative Manipulation for Scientific Data Analysis." In: *Proceedings of the Symposium on Virtual Reality Software and Technology*. VRST'12. Toronto, Ontario, Canada: ACM, 2012, pp. 129–136. DOI: `10.1145/2407336.2407361`.

[Fle+15b] Cédric Fleury, Nicolas Férey, Jean-Marc Vézien, and Patrick Bourdot. "Remote collaboration across heterogeneous large interactive spaces." In: *Workshop on Collaborative Virtual Environments (3DCVE)*

at the *IEEE Conference on Virtual Reality*. 2015, pp. 9–10. DOI: 10.1109/3DCVE.2015.7153591.

[Fle+14] **Cédric Fleury, Tiberiu Popa, Tat Jen Cham, and Henry Fuchs. "Merging Live and pre-Captured Data to support Full 3D Head Reconstruction for Telepresence." In: *Proceedings of the Conference of the European Association for Computer Graphics*. EG'14. Strasbourg, France, 2014, pp. 9–12. DOI: 10.2312/egsh.20141002.**

[FM21] Guo Freeman and Divine Maloney. "Body, Avatar, and Me: The Presentation and Perception of Self in Social Virtual Reality." In: *Proceedings of the ACM on Human-Computer Interaction* 4.CSCW3 (2021). DOI: 10.1145/3432938.

[Fri+21] Sebastian J. Friston, Ben J. Congdon, David Swapp, Lisa Izzouzi, Klara Brandstätter, Daniel Archer, Otto Olkkonen, Felix Johannes Thiel, and Anthony Steed. "Ubiq: A System to Build Flexible Social Virtual Reality Experiences." In: *Proceedings of the Symposium on Virtual Reality Software and Technology*. VRST '21. Osaka, Japan: ACM, 2021. DOI: 10.1145/3489849.3489871.

[FSB14] Henry Fuchs, Andrei State, and Jean-Charles Bazin. "Immersive 3D Telepresence." In: *Computer* 47.7 (2014), pp. 46–52. DOI: 10.1109/MC.2014.185.

[Fuc+96] Henry Fuchs, Andrei State, Etta D. Pisano, William F. Garrett, Gentaro Hirota, Mark A. Livingston, Mary C. Whitton, and Stephen M. Pizer. "Towards Performing Ultrasound-Guided Needle Biopsies from within a Head-Mounted Display." In: *Proceedings of the International Conference on Visualization in Biomedical Computing*. VBC '96. Berlin, Heidelberg: Springer-Verlag, 1996, pp. 591–600. DOI: 10.1007/BFb0047002.

[FMG11] Philippe Fuchs, Guillaume Moreau, and Pascal Guitton. *Virtual Reality: Concepts and Technologies*. 1st. CRC Press, 2011. DOI: 10.1201/b11612.

[Fus+04] Susan R. Fussell, Leslie D. Setlock, Jie Yang, Jiazhi Ou, Elizabeth Mauer, and Adam D.I. Kramer. "Gestures Over Video Streams to Support Remote Collaboration on Physical Tasks." In: *Human–Computer Interaction* 19.3 (2004), pp. 273–309. DOI: 10.1207/s15327051hci1903_3.

[Fyf+18] **Lawrence Fyfe, Olivier Gladin, Cédric Fleury, and Michel Beaudouin-Lafon. "Combining Web Audio Streaming, Motion Capture, and Binaural Audio in a Telepresence System." In: *Proceedings of the International Web Audio conference*. WAC'18. Berlin, Germany, 2018. URL: https://hal.science/hal-01957843.**

[Gau+14] Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. "World-Stabilized Annotations and Virtual Scene Navigation for Remote Collaboration." In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST '14. Honolulu, Hawaii, USA: ACM, 2014, pp. 449–459. DOI: 10.1145/2642918.2647372.

[GBR12] Florian Geyer, Jochen Budzinski, and Harald Reiterer. “IdeaVis: A Hybrid Workspace and Interactive Visualization for Paper-Based Collaborative Sketching Sessions.” In: *Proceedings of the Nordic Conference on Human-Computer Interaction: Making Sense Through Design*. NordiCHI ’12. Copenhagen, Denmark: ACM, 2012, pp. 331–340. DOI: 10.1145/2399016.2399069.

[Gig+14] Dominik Giger, Jean-Charles Bazin, Claudia Kuster, Tiberiu Popa, and Markus Gross. “Gaze correction with a single webcam.” In: *Proceedings of IEEE International Conference on Multimedia and Expo (ICME)*. 2014, pp. 1–6. DOI: 10.1109/ICME.2014.6890306.

[Got+18] Daniel Gotsch, Xujing Zhang, Timothy Merritt, and Roel Vertegaal. “TeleHuman2: A Cylindrical Light Field Teleconferencing System for Life-Size 3D Human Telepresence.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’18. Montreal QC, Canada: ACM, 2018, pp. 1–10. DOI: 10.1145/3173574.3174096.

[Grø+21] Jens Emil Grønbæk, Banu Saatçi, Carla F. Griggio, and Clemens Nylandsted Klokmose. “MirrorBlender: Supporting Hybrid Meetings with a Malleable Video-Conferencing System.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’21. Yokohama, Japan: ACM, 2021. DOI: 10.1145/3411764.3445698.

[Gro+01] Tovi Grossman, Ravin Balakrishnan, Gordon Kurtenbach, George Fitzmaurice, Azam Khan, and Bill Buxton. “Interaction Techniques for 3D Modeling on Large Displays.” In: *Proceedings of the Symposium on Interactive 3D Graphics*. I3D ’01. ACM, 2001, pp. 17–23. DOI: 10.1145/364338.364341.

[Gru01] Jonathan Grudin. “Partitioning Digital Worlds: Focal and Peripheral Awareness in Multiple Monitor Use.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’01. Seattle, Washington, USA: ACM, 2001, pp. 458–465. DOI: 10.1145/365024.365312.

[Gug+17] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. “ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’17. Denver, Colorado, USA: ACM, 2017, pp. 4021–4033. DOI: 10.1145/3025453.3025683.

[Gui+16] Hao Gui, Rong Zheng, Chao Ma, Hao Fan, and Liya Xu. “An Architecture for Healthcare Big Data Management and Analysis.” In: *Health Information Science (HIS)*. Vol. 10038. Lecture Notes in Computer Science. Springer, 2016, pp. 154–160. DOI: 10.1007/978-3-319-48335-1_17.

[Hab+18] Jacob Habgood, David Moore, David Wilson, and Sergio Alapont. “Rapid, Continuous Movement Between Nodes as an Accessible Virtual Reality Locomotion Technique.” In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2018, pp. 371–378. DOI: 10.1109/VR.2018.8446130.

[HCC07] Mark Hancock, Sheelagh Carpendale, and Andy Cockburn. "Shallow-depth 3D Interaction: Design and Evaluation of One-, Two- and Three-touch Techniques." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '07. San Jose, California, USA: ACM, 2007, pp. 1147–1156. DOI: 10.1145/1240624.1240798.

[HCC09] Mark Hancock, Thomas ten Cate, and Sheelagh Carpendale. "Sticky Tools: Full 6DOF Force-based Interaction for Multi-touch Tables." In: *Proceedings of the International Conference on Interactive Tabletops and Surfaces*. ITS '09. Banff, Alberta, Canada: ACM, 2009, pp. 133–140. DOI: 10.1145/1731903.1731930.

[Han97] Chris Hand. "A Survey of 3D Interaction Techniques." In: *Computer Graphics Forum* 16.5 (1997), pp. 269–281. DOI: 10.1111/1467-8659.00194.

[HC02] Denise Y.P. Henriques and John D. Crawford. "Role of Eye, Head, and Shoulder Geometry in the Planning of Accurate Arm Movements." In: *Journal of Neurophysiology* 87.4 (2002), pp. 1677–1685. DOI: 10.1152/jn.00509.2001.

[HYS16] Keita Higuch, Ryo Yonetani, and Yoichi Sato. "Can Eye Help You? Effects of Visualizing Eye Fixations on Remote Collaboration Scenarios for Physical Tasks." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '16. San Jose, California, USA: ACM, 2016, pp. 5180–5190. DOI: 10.1145/2858036.2858438.

[Hig+15] Keita Higuchi, Yinpeng Chen, Philip A. Chou, Zhengyou Zhang, and Zicheng Liu. "ImmerseBoard: Immersive Telepresence Experience Using a Digital Whiteboard." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: ACM, 2015, pp. 2383–2392. DOI: 10.1145/2702123.2702160.

[Hof98] Hunter G. Hoffman. "Physically touching virtual objects using tactile augmentation enhances the realism of virtual environments." In: *Proceedings of the IEEE Virtual Reality Annual International Symposium*. 1998, pp. 59–63. DOI: 10.1109/VRAIS.1998.658423.

[HS92] Jim Hollan and Scott Stornetta. "Beyond Being There." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '92. Monterey, California, USA: ACM, 1992, pp. 119–125. DOI: 10.1145/142750.142769.

[HRB97] Michael P. Hollier, Andrew N. Rimell, and D. Burraston. "Spatial audio technology for telepresence." In: *BT Technology Journal* 15 (1997), pp. 33–41. DOI: 10.1023/A:1018675327815.

[Hor+18] Tom Horak, Sriram Karthik Badam, Niklas Elmqvist, and Raimund Dachselt. "When David Meets Goliath: Combining Smartwatches with a Large Vertical Display for Visual Data Exploration." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: ACM, 2018, pp. 1–13. DOI: 10.1145/3173574.3173593.

[IIH07] Takeo Igarashi, Takeo Igarashi, and John F. Hughes. “A Suggestive Interface for 3D Drawing.” In: *ACM SIGGRAPH 2007 Courses*. SIGGRAPH ’07. San Diego, California: ACM, 2007. DOI: 10.1145/1281500.1281531.

[Irl+16] Andrew Irlitti, Ross T. Smith, Stewart Von Itzstein, Mark Billinghurst, and Bruce H. Thomas. “Challenges for Asynchronous Collaboration in Augmented Reality.” In: *Adjunct proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*. 2016, pp. 31–35. DOI: 10.1109/ISMAR-Adjunct.2016.0032.

[IT93] Ellen A. Isaacs and John C. Tang. “What Video Can and Can’T Do for Collaboration: A Case Study.” In: *Proceedings of the International Conference on Multimedia*. MULTIMEDIA ’93. Anaheim, California, USA: ACM, 1993, pp. 199–206. DOI: 10.1145/166266.166289.

[IK92] Hiroshi Ishii and Minoru Kobayashi. “ClearBoard: A Seamless Medium for Shared Drawing and Conversation with Eye Contact.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’92. Monterey, California, USA: ACM, 1992, pp. 525–532. DOI: 10.1145/142750.142977.

[Isr+09] Johann Israel, Eva Wiese, Magdalena Mateescu, Christian Zöllner, and Rainer Stark. “Investigating three-dimensional sketching for early conceptual design—Results from expert discussions and user studies.” In: *Computers & Graphics* 33.4 (2009), pp. 462–473. DOI: 10.1016/j.cag.2009.05.005.

[Iza+11] Shahram Izadi et al. “KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Camera.” In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST ’11. ACM, 2011, pp. 559–568. DOI: 10.1145/2047196.2047270.

[JSK12] Bret Jackson, David Schroeder, and Daniel F. Keefe. “Nailing Down Multi-Touch: Anchored Above the Surface Interaction for 3D Modeling and Navigation.” In: *Proceedings of the Graphics Interface 2012*. GI ’12. Canadian Information Processing Society, 2012, pp. 181–184. URL: https://dl.acm.org/doi/10.5555/2305276.2305306.

[JFH94] Richard H. Jacoby, Mark Ferneau, and Jim Humphries. “Gestural interaction in a virtual environment.” In: *Stereoscopic Displays and Virtual Reality Systems*. Ed. by Scott S. Fisher, John O. Merritt, and Mark T. Bolas. Vol. 2177. International Society for Optics and Photonics. SPIE, 1994, pp. 355–364. DOI: 10.1117/12.173892.

[JH14] Mikkel R. Jakobsen and Kasper Hornbæk. “Up Close and Personal: Collaborative Work on a High-Resolution Multitouch Wall Display.” In: *ACM Transactions on Computer-Human Interaction* 21.2 (2014). DOI: 10.1145/2576099.

[JH16] Mikkel R. Jakobsen and Kasper Hornbæk. “Negotiating for Space? Collaborative Work Using a Wall Display with Mouse and Touch Input.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’16. San Jose, California, USA: ACM, 2016, pp. 2050–2061. DOI: 10.1145/2858036.2858158.

[JBC23] Raphaël James, Anastasia Bezerianos, and Olivier Chapuis. "Evaluating the Extension of Wall Displays with AR for Collaborative Work." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '23. Hamburg, Germany: ACM, 2023. DOI: 10.1145/3544548.3580752.

[Jan+20] Pascal Jansen, Fabian Fischbach, Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. "ShARe: Enabling Co-Located Asymmetric Multi-User Interaction for Augmented Reality Head-Mounted Displays." In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST '20. Virtual Event, USA: ACM, 2020, pp. 459–471. DOI: 10.1145/3379337.3415843.

[JSH19] Yvonne Jansen, Jonas Schjerlund, and Kasper Hornbæk. "Effects of Locomotion and Visual Overview on Spatial Memory When Interacting with Wall Displays." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: ACM, 2019, pp. 1–12. DOI: 10.1145/3290605.3300521.

[Jon+21] Brennan Jones, Yaying Zhang, Priscilla N. Y. Wong, and Sean Rintel. "Belonging There: VROOM-Ing into the Uncanny Valley of XR Telepresence." In: *Proceedings of the ACM on Human-Computer Interaction* 5.CSCW1 (2021). DOI: 10.1145/3449133.

[Kam+18] Adhitya Kamakshidasan, José Galaz, Rodrigo Cienfuegos, Antoine Rousseau, and Emmanuel Pietriga. "Comparative Visualization of Deep Water Asteroid Impacts on Ultra-high-resolution Wall Displays with Seawall." In: *Proceedings of the IEEE Scientific Visualization Conference*. SciVis'18. 2018, pp. 142–145. DOI: 10.1109/SciVis.2018.8823616.

[KRF11] Anette von Kapri, Tobias Rick, and Steven Feiner. "Comparing steering-based travel techniques for search tasks in a CAVE." In: *Proceedings of the IEEE Virtual Reality Conference (VR)*. 2011, pp. 91–94. DOI: 10.1109/VR.2011.5759443.

[Kaz+17] Rubaiat Habib Kazi, Tovi Grossman, Hyunmin Cheong, Ali Hashemi, and George Fitzmaurice. "DreamSketch: Early Stage 3D Design Explorations with Sketching and Generative Design." In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST '17. Québec City, QC, Canada: ACM, 2017, pp. 401–414. DOI: 10.1145/3126594.3126662.

[KZL07] Daniel F. Keefe, Robert C. Zeleznik, and David H. Laidlaw. "Drawing on Air: Input Techniques for Controlled 3D Line Illustration." In: *IEEE Transactions on Visualization and Computer Graphics* 13.5 (2007), pp. 1067–1081. DOI: 10.1109/TVCG.2007.1060.

[KD07] Fakheredine Keyrouz and Klaus Diepold. "Binaural Source Localization and Spatial Audio Reproduction for Telepresence Applications." In: *Presence: Teleoperators and Virtual Environments* 16.5 (2007), pp. 509–522. DOI: 10.1162/pres.16.5.509.

[Kha+05] Azam Khan, Justin Matejka, George Fitzmaurice, and Gordon Kurtenbach. “Spotlight: Directing Users’ Attention on Large Displays.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’05. Portland, Oregon, USA: ACM, 2005, pp. 791–798. DOI: `10.1145/1054972.1055082`.

[Kis+17] Ulrike Kister, Konstantin Klamka, Christian Tominski, and Raimund Dachselt. “GraSp: Combining Spatially-Aware Mobile Devices and a Display Wall for Graph Visualization and Interaction.” In: *Computer Graphics Forum* 36.3 (2017), pp. 503–514. DOI: `10.1111/cgf.13206`.

[Kiy+02] Kiyoshi Kiyokawa, Mark Billinghurst, Sean Hayes, Anoop Gupta, Yuki Sannohe, and Hirokazu Kato. “Communication behaviors of co-located users in collaborative AR interfaces.” In: *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR)*. 2002, pp. 139–148. DOI: `10.1109/ISMAR.2002.1115083`.

[KMI19] Anya Kolesnichenko, Joshua McVeigh-Schultz, and Katherine Isbister. “Understanding Emerging Design Practices for Avatar Systems in the Commercial Social VR Ecology.” In: *Proceedings of the Designing Interactive Systems Conference*. DIS ’19. San Diego, CA, USA: ACM, 2019, pp. 241–252. DOI: `10.1145/3322276.3322352`.

[KS04] Vladislav Kraevoy and Alla Sheffer. “Cross-parameterization and Compatible Remeshing of 3D Models.” In: *ACM Transactions on Graphics (siggraph’04)* 23.3 (2004), pp. 861–869. DOI: `10.1145/1015706.1015811`.

[KK06] Martin Kuechler and Andreas Kunz. “HoloPort - A Device for Simultaneous Video and Data Conferencing Featuring Gaze Awareness.” In: *Proceedings of IEEE Virtual Reality Conference (VR)*. 2006, pp. 81–88. DOI: `10.1109/VR.2006.71`.

[Kul+11] Alexander Kulik, André Kunert, Stephan Beck, Roman Reichel, Roland Blach, Armin Zink, and Bernd Froehlich. “C1x6: A Stereoscopic Six-User Display for Co-Located Collaboration in Shared Virtual Environments.” In: *ACM Transactions on Graphics* 30.6 (2011), pp. 1–12. DOI: `10.1145/2070781.2024222`.

[Kun+14] André Kunert, Alexander Kulik, Stephan Beck, and Bernd Froehlich. “Photoportals: Shared References in Space and Time.” In: *Proceedings of the Conference on Computer Supported Cooperative Work & Social Computing*. CSCW ’14. Baltimore, Maryland, USA: ACM, 2014, pp. 1388–1399. DOI: `10.1145/2531602.2531727`.

[LaV+17] Joseph J. LaViola, Ernst Kruijff, Ryan P. McMahan, Doug A. Bowman, and Ivan Poupyrev. *3D User Interfaces: Theory and Practice, 2nd edition*. Addison-Wesley usability and HCI series. Addison Wesley Longman Publishing Co., Inc., 2017. ISBN: 9780134034324.

[Le+19] Khanh-Duy Le, Ignacio Avellino, Cédric Fleury, Morten Fjeld, and Andreas M. Kunz. “GazeLens: Guiding Attention to Improve Gaze Interpretation in Hub-Satellite Collaboration.” In: *Proceedings of the IFIP TC13 International Conference On Human-Computer Interaction (INTERACT)*. Vol. 11747. Lecture Notes in Computer Science. Paphos,

Cyprus: Springer, 2019, pp. 282–303. DOI: 10.1007/978-3-030-29384-0_18.

[Léc+23] Aurélien Léchappé, Aurélien Milliat, Cédric Fleury, Mathieu Chollet, and Cédric Dumas. "Characterization of Collaboration in a Virtual Environment with Gaze and Speech Signals." In: *Companion Publication of the International Conference on Multimodal Interaction*. ICMI '23 Companion. Paris, France: ACM, 2023, pp. 21–25. DOI: 10.1145/3610661.3617149.

[LMR14] Nicolas H. Lehment, Daniel Merget, and Gerhard Rigoll. "Creating automatically aligned consensus realities for AR videoconferencing." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2014, pp. 201–206. DOI: 10.1109/ISMAR.2014.6948428.

[LD97] Valerie D. Lehner and Thomas A. DeFanti. "Distributed virtual reality: supporting remote collaboration in vehicle design." In: *IEEE Computer Graphics and Applications* 17.2 (1997), pp. 13–17. DOI: 10.1109/38.574654.

[Lei+96] Jason Leigh, Andrew E. Johnson, Christina A. Vasilakis, and Thomas A. DeFanti. "Multi-perspective collaborative design in persistent networked virtual environments." In: *Proceedings of the IEEE Virtual Reality Annual International Symposium*. 1996, pp. 253–260. DOI: 10.1109/VRAIS.1996.490535.

[Li+13] Hao Li, Jihun Yu, Yuting Ye, and Chris Bregler. "Realtime Facial Animation with On-the-fly Correctives." In: *ACM Transactions on Graphics (siggraph'13)* 32.4 (2013), 42:1–10. DOI: 10.1145/2461912.2462019.

[Lis+15] Lars Lischke, Sven Mayer, Katrin Wolf, Niels Henze, Albrecht Schmidt, Svenja Leifert, and Harald Reiterer. "Using Space: Effect of Display Size on Users' Search Performance." In: *Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems*. CHI EA '15. Seoul, Republic of Korea: ACM, 2015, pp. 1845–1850. DOI: 10.1145/2702613.2732845.

[Liu15] Can Liu. "Embodied Interaction for Data Manipulation Tasks on Wall-sized Displays." PhD thesis. Université Paris-Saclay, 2015. URL: https://theses.hal.science/tel-01264670v2.

[Liu+16] Can Liu, Olivier Chapuis, Michel Beaudouin-Lafon, and Eric Lecolinet. "Shared Interaction on a Wall-Sized Display in a Data Manipulation Task." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '16. San Jose, California, USA: ACM, 2016, pp. 2075–2086. DOI: 10.1145/2858036.2858039.

[Liu+17] Can Liu, Olivier Chapuis, Michel Beaudouin-Lafon, and Eric Lecolinet. "CoReach: Cooperative Gestures for Data Manipulation on Wall-Sized Displays." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: ACM, 2017, pp. 6730–6741. DOI: 10.1145/3025453.3025594.

[Liu+14] Can Liu, Olivier Chapuis, Michel Beaudouin-Lafon, Eric Lecolinet, and Wendy E. Mackay. “Effects of Display Size and Navigation Type on a Classification Task.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’14. Toronto, Ontario, Canada: ACM, 2014, pp. 4147–4156. DOI: `10.1145/2556288.2557020`.

[Liu+18] James Liu, Hirav Parekh, Majed Al-Zayer, and Eelke Folmer. “Increasing Walking in VR Using Redirected Teleportation.” In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST ’18. Berlin, Germany: ACM, 2018, pp. 521–529. DOI: `10.1145/3242587.3242601`.

[LBC04] Julian Looser, Mark Billinghurst, and Andy Cockburn. “Through the Looking Glass: The Use of Lenses as an Interface Tool for Augmented Reality Interfaces.” In: *Proceedings of the International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia*. GRAPHITE ’04. Singapore: ACM, 2004, pp. 204–211. DOI: `10.1145/988834.988870`.

[Lop+16] David Lopez, Lora Oehlberg, Candemir Doger, and Tobias Isenberg. “Towards An Understanding of Mobile Touch Navigation in a Stereoscopic Viewing Environment for 3D Data Exploration.” In: *IEEE Transactions on Visualization and Computer Graphics* 22.5 (2016), pp. 1616–1629. DOI: `10.1109/TVCG.2015.2440233`.

[LF16] Jean-Baptiste Louvet and Cédric Fleury. “Combining Bimanual Interaction and Teleportation for 3D Manipulation on Multi-Touch Wall-Sized Displays.” In: *Proceedings of the Symposium on Virtual Reality Software and Technology*. VRST’16. Munich, Germany: ACM, 2016, pp. 283–292. DOI: `10.1145/2993369.2993390`.

[Luf+15] Paul K. Luff, Naomi Yamashita, Hideaki Kuzuoka, and Christian Heath. “Flexible Ecologies And Incongruent Locations.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’15. Seoul, Republic of Korea: ACM, 2015, pp. 877–886. DOI: `10.1145/2702123.2702286`.

[Ma+21] Shugao Ma, Tomas Simon, Jason Saragih, Dawei Wang, Yuecheng Li, Fernando De La Torre, and Yaser Sheikh. “Pixel Codec Avatars.” In: *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE Computer Society, 2021, pp. 64–73. DOI: `10.1109/CVPR46437.2021.00013`.

[Mac99] Wendy E. Mackay. “Media Spaces: Environments for multimedia interaction.” In: *Computer-Supported Cooperative Work*. Ed. by Michel Beaudouin-Lafon. Trends in Software Series. Chichester: Wiley and Sons, 1999, pp. 55–82. ISBN: 0-471-96736-X.

[Mah+10] Morad Mahdjoub, Davy Monticolo, Samuel Gomes, and Jean-Claude Sagot. “A Collaborative Design for Usability Approach Supported by Virtual Reality and a Multi-Agent System Embedded in a PLM Environment.” In: *Computer-Aided Design* 42.5 (2010), pp. 402–413. DOI: `10.1016/j.cad.2009.02.009`.

[Mah+19] Tahir Mahmood, Willis Fulmer, Neelesh Mungoli, Jian Huang, and Aidong Lu. "Improving Information Sharing and Collaborative Analysis for Remote GeoSpatial Visualization Using Mixed Reality." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2019, pp. 236–247. DOI: 10.1109/ISMAR.2019.00021.

[MHG12] Nicolai Marquardt, Ken Hinckley, and Saul Greenberg. "Cross-Device Interaction via Micro-Mobility and f-Formations." In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST '12. Cambridge, Massachusetts, USA: ACM, 2012, pp. 13–22. DOI: 10.1145/2380116.2380121.

[Mar+17] Pierre Martin, Stéphane Masfrand, Yujiro Okuya, and Patrick Bourdot. "A VR-CAD Data Model for Immersive Design." In: *Augmented Reality, Virtual Reality, and Computer Graphics*. Cham: Springer International Publishing, 2017, pp. 222–241. DOI: 10.1007/978-3-319-60922-5_17.

[MCG10a] Anthony Martinet, Gery Casiez, and Laurent Grisoni. "The Design and Evaluation of 3D Positioning Techniques for Multi-touch Displays." In: *Proceedings of the IEEE Symposium on 3D User Interfaces*. 3DUI '10. Washington, DC, USA: IEEE Computer Society, 2010, pp. 115–118. DOI: 10.1109/3DUI.2010.5444709.

[MCG10b] Anthony Martinet, Géry Casiez, and Laurent Grisoni. "The Effect of DOF Separation in 3D Manipulation Tasks with Multi-touch Displays." In: *Proceedings of the Symposium on Virtual Reality Software and Technology*. VRST '10. Hong Kong: ACM, 2010, pp. 111–118. DOI: 10.1145/1889863.1889888.

[Mil+95] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. "Augmented reality: a class of displays on the reality-virtuality continuum." In: *Telemanipulator and Telepresence Technologies*. Ed. by Hari Das. Vol. 2351. International Society for Optics and Photonics. SPIE, 1995, pp. 282–292. DOI: 10.1117/12.197321.

[Mil+93] Paul Milgram, Shumin Zhai, David Drascic, and Julius Grodski. "Applications of augmented reality for human-robot communication." In: *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '93)*. Vol. 3. 1993, 1467–1472 vol.3. DOI: 10.1109/IROS.1993.583833.

[Min+20] Dae-Hong Min, Dong-Yong Lee, Yong-Hun Cho, and In-Kwon Lee. "Shaking Hands in Virtual Space: Recovery in Redirected Walking for Direct Interaction between Two Users." In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2020, pp. 164–173. DOI: 10.1109/VR46266.2020.00035.

[Min95] Mark R. Mine. *Virtual Environment Interaction Techniques*. Tech. rep. University of North Carolina at Chapel Hill, 1995.

[Moh+20] Peter Mohr, Shohei Mori, Tobias Langlotz, Bruce H. Thomas, Dieter Schmalstieg, and Denis Kalkofen. "Mixed Reality Light Fields for Interactive Remote Assistance." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '20. Honolulu, HI, USA: ACM, 2020, pp. 1–12. DOI: 10.1145/3313831.3376289.

[MG02] Andrew F. Monk and Caroline Gale. "A Look Is Worth a Thousand Words: Full Gaze Awareness in Video-Mediated Conversation." In: *Discourse Processes* 33.3 (2002), pp. 257–278. DOI: 10.1207/S15326950DP3303_4.

[MMK12] Masahiro Mori, Karl F. MacDorman, and Norri Kageki. "The Uncanny Valley [From the Field]." In: *IEEE Robotics & Automation Magazine* 19.2 (2012), pp. 98–100. DOI: 10.1109/MRA.2012.2192811.

[MSH04] Tariq Mujber, Tamas Szecsi, and Saleem Hashmi. "Virtual reality applications in manufacturing process simulation." In: *Journal of Materials Processing Technology* 155-156 (2004), pp. 1834–1838. DOI: 10.1016/j.jmatprotec.2004.04.401.

[Nan+15] Mathieu Nancel, Emmanuel Pietriga, Olivier Chapuis, and Michel Beaudouin-Lafon. "Mid-Air Pointing on Ultra-Walls." In: *ACM Transactions on Computer-Human Interaction* 22.5 (2015). DOI: 10.1145/2766448.

[NC05] David Nguyen and John Canny. "MultiView: Spatially Faithful Group Video Conferencing." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '05. Portland, Oregon, USA: ACM, 2005, pp. 799–808. DOI: 10.1145/1054972.1055084.

[NM97] Haruo Noma and Tsutomu Miyasato. "Cooperative Object Manipulation in Virtual Space Using Virtual Physics." In: *Proceedings of the ASME International Mechanical Engineering Congress and Exposition*. Vol. Dynamic Systems and Control. Nov. 1997, pp. 101–106. DOI: 10.1115/IMECE1997-0383.

[Obs22] Rubin Observatory. *Data Management*. https://www.lsst.org/about/dm, accessed: 2022-03-23. 2022.

[OF09] Ohan Oda and Steven Feiner. "Interference avoidance in multi-user hand-held augmented reality." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*. 2009, pp. 13–22. DOI: 10.1109/ISMAR.2009.5336507.

[Oku+20] Yujiro Okuya, Olivier Gladin, Nicolas Ladévèze, Cédric Fleury, and Patrick Bourdot. "Investigating Collaborative Exploration of Design Alternatives on a Wall-Sized Display." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI'20. Virtual Event: ACM, 2020, pp. 1–12. DOI: 10.1145/3313831.3376736.

[Oku+18a] Yujiro Okuya, Nicolas Ladévèze, Cédric Fleury, and Patrick Bourdot. "ShapeGuide: Shape-Based 3D Interaction for Parameter Modification of Native CAD Data." In: *Frontiers in Robotics and AI* 5 (2018). DOI: 10.3389/frobt.2018.00118.

[Oku+18b] Yujiro Okuya, Nicolas Ladévèze, Olivier Gladin, Cédric Fleury, and Patrick Bourdot. "Distributed Architecture for Remote Collaborative Modification of Parametric CAD Data." In: *Workshop on Collaborative Virtual Environments (3DCVE) at the IEEE Conference on Virtual Reality*. Reutlingen, Germany, 2018, pp. 1–4. DOI: 10.1109/3DCVE.2018.8637112.

[Oku+21] Yujiro Okuya, Nicolas Ladévèze, Olivier Gladin, Cédric Fleury, and Patrick Bourdot. "Collaborative VR-CAD for Industrial Product Design: CAD Parameter Modification with 3D Interaction on Heterogeneous Immersive Platforms." In: *Manufacturing in the Era of 4th Industrial Revolution*. Ed. by Monica Bordegoni, Satyandra K. Gupta, and James Ritchie. Vol. 3. World Scientific, 2021. Chap. 2, pp. 17–47. DOI: 10.1142/9789811222863_0002.

[Ort+16] Sergio Orts-Escolano et al. "Holoportation: Virtual 3D Teleportation in Real-Time." In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST '16. Tokyo, Japan: ACM, 2016, pp. 741–754. DOI: 10.1145/2984511.2984517.

[Ots16] Kazuhiro Otsuka. "MMSpace: Kinetically-augmented telepresence for small group-to-group conversations." In: *Proceedings of IEEE Virtual Reality (VR)*. 2016, pp. 19–28. DOI: 10.1109/VR.2016.7504684.

[PS14] Ye Pan and Anthony Steed. "A Gaze-Preserving Situated Multiview Telepresence System." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada: ACM, 2014, pp. 2173–2176. DOI: 10.1145/2556288.2557320.

[PL06] Fred Pighin and John P. Lewis. "Performance-Driven Facial Animation." In: *ACM Siggraph Courses*. 2006. DOI: 10.1145/1185657.1185841. URL: http://scribblethink.org/Courses/PDFA/2006_COURSE30.pdf.

[PBF08] Marcio S. Pinho, Doug A. Bowman, and Carla M. Dal Sasso Freitas. "Cooperative object manipulation in collaborative virtual environments." In: *Journal of the Brazilian Computer Society* 14 (2008), pp. 53–67. DOI: 10.1007/BF03192559.

[PBF02] Márcio S. Pinho, Doug A. Bowman, and Carla M.D.S. Freitas. "Cooperative Object Manipulation in Immersive Virtual Environments: Framework and Techniques." In: *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. VRST '02. Hong Kong, China: ACM, 2002, pp. 171–178. DOI: 10.1145/585740.585769.

[Piu+19] Thammathip Piumsomboon, Gun A. Lee, Andrew Irlitti, Barrett Ens, Bruce H. Thomas, and Mark Billinghurst. "On the Shoulder of the Giant: A Multi-Scale Mixed Reality Collaboration with 360 Video Sharing and Tangible Interaction." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: ACM, 2019, pp. 1–17. DOI: 10.1145/3290605.3300458.

[PBN22] Carole Plasson, Renaud Blanch, and Laurence Nigay. "Selection Techniques for 3D Extended Desktop Workstation with AR HMD." In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2022, pp. 460–469. DOI: 10.1109/ISMAR55827.2022.00062.

[Pou+96] Ivan Poupyrev, Mark Billinghurst, Suzanne Weghorst, and Tadao Ichikawa. "The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR." In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST '96. Seattle, Washington, USA: ACM, 1996, pp. 79–80. DOI: 10.1145/237091.237102.

[PBC17] Arnaud Prouzeau, Anastasia Bezerianos, and Olivier Chapuis. "Evaluating Multi-User Selection for Exploring Graph Topology on Wall-Displays." In: *IEEE Transactions on Visualization and Computer Graphics* 23.8 (2017), pp. 1936–1951. DOI: 10.1109/TVCG.2016.2592906.

[PBC18] Arnaud Prouzeau, Anastasia Bezerianos, and Olivier Chapuis. "Awareness Techniques to Aid Transitions between Personal and Shared Workspaces in Multi-Display Environments." In: *Proceedings of the International Conference on Interactive Surfaces and Spaces*. ISS'18. Tokyo, Japan: ACM, 2018, pp. 291–304. DOI: 10.1145/3279778.3279780.

[RR14] Wullianallur Raghupathi and Viju Raghupathi. "Big data analytics in healthcare: promise and potential." In: *Health Information Science and Systems* 2.3 (2014). DOI: 10.1186/2047-2501-2-3.

[Ran+01] Dirk Rantzau, Franz Maurer, C. Mayer, Robin Löffler, Oliver Riedel, Holger R. Scharm, and D. Banek. "The integration of immersive Virtual Reality applications into Catia V5." In: *Immersive Projection Technology and Virtual Environments*. Vienna: Springer Vienna, 2001, pp. 93–102. DOI: 10.1007/978-3-7091-6221-7_10.

[Rap+09] Alberto Raposo, Ismael Santos, Luciano Soares, Gustavo Wagner, Eduardo Corseuil, and Marcelo Gattass. "Environ: Integrating VR and CAD in Engineering Projects." In: *IEEE Computer Graphics and Applications* 29.6 (2009), pp. 91–95. DOI: 10.1109/MCG.2009.118.

[RKW01] Sharif Razzaque, Zachariah Kohn, and Mary C. Whitton. "Redirected Walking." In: *Proceedings of the Eurographics conference - Short Presentations*. Eurographics Association, 2001. DOI: 10.2312/egs.20011036.

[RS02] Holger Regenbrecht and Thomas Schubert. "Real and illusory interactions enhance presence in virtual environments." In: *Presence: Teleoperators & Virtual Environments* 11.4 (2002), pp. 425–434. DOI: 10.1162/105474602760204318.

[RDH09] Jason L. Reisman, Philip L. Davidson, and Jefferson Y. Han. "A Screen-space Formulation for 2D and 3D Direct Manipulation." In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST '09. Victoria, BC, Canada: ACM, 2009, pp. 69–78. DOI: 10.1145/1622176.1622190.

[Rek96] Jun Rekimoto. "Transvision: A hand-held augmented reality system for collaborative design." In: *Proceedings of Virtual Systems and Multi-Media (VSMM '96)*. Jan. 1996.

[Rie+06] Kai Riege, Thorsten Holtkamper, Gerold Wesche, and Bernd Froehlich. "The Bent Pick Ray: An Extended Pointing Technique for Multi-User Interaction." In: *Proceedings of IEEE Symposium on 3D User Interfaces (3DUI'06)*. 2006, pp. 62–65. DOI: 10.1109/VR.2006.127.

[Rim22] Meghan Rimol. *Gartner Survey Reveals a 44% Rise in Workers' Use of Collaboration Tools Since 2019*. https://www.gartner.com/en/newsroom/press-releases/2021-08-23-gartner-survey-reveals-44-percent-rise-in-workers-use-of-collaboration-tools-since-2019, accessed: 2022-03-23. 2022.

[RH17] Joan Sol Roo and Martin Hachet. “One Reality: Augmenting How the Physical World is Experienced by Combining Multiple Mixed Reality Modalities.” In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST ’17. Québec City, QC, Canada: ACM, 2017, pp. 787–795. DOI: 10.1145/3126594.3126638.

[RSJ02] Roy A. Ruddle, Justin C. D. Savage, and Dylan M. Jones. “Symmetric and Asymmetric Action Integration during Cooperative Object Manipulation in Virtual Environments.” In: *ACM Transactions on Computer-Human Interaction* 9.4 (2002), pp. 285–308. DOI: 10.1145/586081.586084.

[Rud+16] Roy A. Ruddle, Rhys G. Thomas, Rebecca Randell, Philip Quirke, and Darren Treanor. “The Design and Evaluation of Interfaces for Navigating Gigapixel Images in Digital Pathology.” In: *ACM Transactions on Computer-Human Interaction* 23.1 (2016). DOI: 10.1145/2834117.

[SJF09] Holger Salzmann, Jan Jacobs, and Bernd Froehlich. “Collaborative Interaction in Co-Located Two-User Scenarios.” In: *Proceedings of the Joint Virtual Reality Eurographics Conference on Virtual Environments*. JVRC’09. Lyon, France: Eurographics Association, 2009, pp. 85–92. DOI: 10.2312/EGVE/JVRC09/085-092.

[Sch03] Thomas W. Schubert. “The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness.” In: *Zeitschrift für Medienpsychologie* 15.2 (2003), pp. 69–71. DOI: 10.1026//1617-6383.15.2.69.

[Sch+12] Tobias Schwarz, Simon Butscher, Jens Mueller, and Harald Reiterer. “Content-Aware Navigation for Large Displays in Context of Traffic Control Rooms.” In: *Proceedings of the International Working Conference on Advanced Visual Interfaces*. AVI’12. Capri Island, Italy: ACM, 2012, pp. 249–252. DOI: 10.1145/2254556.2254601.

[SBA92] Abigail Sellen, Bill Buxton, and John Arnott. “Using Spatial Cues to Improve Videoconferencing.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’92. Monterey, California, USA: ACM, 1992, pp. 651–652. DOI: 10.1145/142750.143070.

[Ser+22] Mickael Sereno, Xiyao Wang, Lonni Besançon, Michael J. McGuffin, and Tobias Isenberg. “Collaborative Work in Augmented Reality: A Survey.” In: *IEEE Transactions on Visualization and Computer Graphics* 28.6 (2022), pp. 2530–2549. DOI: 10.1109/TVCG.2020.3032761.

[SVG15] Adalberto L. Simeone, Eduardo Velloso, and Hans Gellersen. “Substitutional Reality: Using the Physical Environment to Design Virtual Reality Experiences.” In: *Proceedings of the Conference on Human Factors in Computing Systems*. CHI ’15. Seoul, Republic of Korea: ACM, 2015, pp. 3307–3316. DOI: 10.1145/2702123.2702389.

[SSW21] Richard Skarbez, Missie Smith, and Mary C. Whitton. “Revisiting Milgram and Kishino’s Reality-Virtuality Continuum.” In: *Frontiers in Virtual Reality* 2 (2021). DOI: 10.3389/frvir.2021.647997.

[SH16] Drew Skillman and Patrick Hackett. *Tilt Brush application, Google Inc.* 2016. URL: https://www.tiltbrush.com/.

[Sra+16] Misha Sra, Sergio Garrido-Jurado, Chris Schmandt, and Pattie Maes. "Procedurally generated virtual reality from 3D reconstructed physical space." In: *Proceedings of the Conference on Virtual Reality Software and Technology*. 2016, pp. 191–200. DOI: 10.1145/2993369.2993372.

[Ste+10] Frank Steinicke, Gerd Bruder, Jason Jerald, Harald Frenz, and Markus Lappe. "Estimation of Detection Thresholds for Redirected Walking Techniques." In: *IEEE Transactions on Visualization and Computer Graphics* 16.1 (2010), pp. 17–27. DOI: 10.1109/TVCG.2009.62.

[SB02] Joachim Stempfle and Petra Badke-Schaub. "Thinking in design teams - an analysis of team communication." In: *Design Studies* 23.5 (2002), pp. 473–496. DOI: 10.1016/S0142-694X(02)00004-2.

[SCP95] Richard Stoakley, Matthew J. Conway, and Randy Pausch. "Virtual Reality on a WIM: Interactive Worlds in Miniature." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '95. Denver, Colorado, USA: ACM Press/Addison-Wesley Publishing Co., 1995, pp. 265–272. DOI: 10.1145/223904.223938.

[Str+99] Norbert A. Streitz, Jörg Geißler, Torsten Holmer, Shin'ichi Konomi, Christian Müller-Tomfelde, Wolfgang Reischl, Petra Rexroth, Peter Seitz, and Ralf Steinmetz. "I-LAND: An Interactive Landscape for Creativity and Innovation." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI'99. Pittsburgh, Pennsylvania, USA: ACM, 1999, pp. 120–127. DOI: 10.1145/302979.303010.

[Sum+12] Evan A. Suma, Zachary Lipps, Samantha Finkelstein, David M. Krum, and Mark Bolas. "Impossible spaces: Maximizing natural walking in virtual environments with self-overlapping architecture." In: *IEEE Transactions on Visualization and Computer Graphics* 18.4 (2012), pp. 555–564. DOI: 10.1109/TVCG.2012.47.

[SWK16] Qi Sun, Li-Yi Wei, and Arie Kaufman. "Mapping virtual and physical reality." In: *ACM Transactions on Graphics (TOG)* 35.4 (2016), pp. 1–12. DOI: 10.1145/2897824.2925883.

[Sza+98] Zsolt Szalavári, Dieter Schmalstieg, Anton Fuhrmann, and Michael Gervautz. ""Studierstube": An environment for collaboration in augmented reality." In: *Virtual Reality* 3.1 (1998), pp. 37–48. DOI: 10.1007/BF01409796.

[TB15] Matthew Tait and Mark Billinghurst. "The Effect of View Independence in a Collaborative AR System." In: *Computer Supported Cooperative Work (CSCW)* 24.6 (2015), pp. 563–589. DOI: 10.1007/s10606-015-9231-8.

[TM90] John C. Tang and Scott L. Minneman. "VideoDraw: A Video Interface for Collaborative Drawing." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '90. Seattle, Washington, USA: ACM, 1990, pp. 313–320. DOI: 10.1145/97243.97302.

[TM91] John C. Tang and Scott L. Minneman. "VideoWhiteboard: Video Shadows to Support Remote Collaboration." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '91. New Orleans, Louisiana, USA: ACM, 1991, pp. 315–322. DOI: `10.1145/108844.108932`.

[Teo+19] Theophilus Teo, Louise Lawrence, Gun A. Lee, Mark Billinghurst, and Matt Adcock. "Mixed Reality Remote Collaboration Combining 360 Video and 3D Reconstruction." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: ACM, 2019, pp. 1–14. DOI: `10.1145/3290605.3300431`.

[Tho+20] Balasaravanan Thoravi Kumaravel, Cuong Nguyen, Stephen DiVerdi, and Bjoern Hartmann. "TransceiVR: Bridging Asymmetrical Communication Between VR Users and External Collaborators." In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST '20. Virtual Event, USA: ACM, 2020, pp. 182–195. DOI: `10.1145/3379337.3415827`.

[Uso+99] Martin Usoh, Kevin Arthur, Mary C. Whitton, Rui Bastos, Anthony Steed, Mel Slater, and Frederick P. Brooks. "Walking > Walking-in-Place > Flying, in Virtual Environments." In: *Proceedings of the Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '99. ACM, 1999, pp. 359–364. DOI: `10.1145/311535.311589`.

[Val+11] Dimitar Valkov, Frank Steinicke, Gerd Bruder, and Klaus Hinrichs. "2D Touching of 3D Stereoscopic Objects." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. Vancouver, BC, Canada: ACM, 2011, pp. 1353–1362. DOI: `10.1145/1978942.1979142`.

[Vei+99] Elizabeth S. Veinott, Judith Olson, Gary M. Olson, and Xiaolan Fu. "Video Helps Remote Work: Speakers Who Need to Negotiate Common Ground Benefit from Seeing Each Other." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '99. Pittsburgh, Pennsylvania, USA: ACM, 1999, pp. 302–309. DOI: `10.1145/302979.303067`.

[Ver99] Roel Vertegaal. "The GAZE Groupware System: Mediating Joint Attention in Multiparty Communication and Collaboration." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '99. Pittsburgh, Pennsylvania, USA: ACM, 1999, pp. 294–301. DOI: `10.1145/302979.303065`.

[Ver+01] Roel Vertegaal, Robert Slagter, Gerrit Van der Veer, and Anton Nijholt. "Eye Gaze Patterns in Conversations: There is More to Conversational Agents than Meets the Eyes." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '01. Seattle, Washington, USA: ACM, 2001, pp. 301–308. DOI: `10.1145/365024.365119`.

[VVV00] Roel Vertegaal, Gerrit Van der Veer, and Harro Vons. "Effects of Gaze on Multiparty Mediated Communication." In: *Proceedings of the Graphics Interface 2000 Conference*. GI '00. Montr'eal, Qu'ebec, Canada, 2000, pp. 95–102. DOI: `10.20380/GI2000.14`.

[Wal+20] Shaun Wallace, Brendan Le, Luis A. Leiva, Aman Haq, Ari Kintisch, Gabrielle Bufrem, Linda Chang, and Jeff Huang. "Sketchy: Drawing Inspiration from the Crowd." In: *Proceedings of the ACM on Human-Computer Interaction* 4.CSCW2 (2020). DOI: 10.1145/3415243.

[Wan+20a] Chiu-Hsuan Wang, Chia-En Tsai, Seraphina Yong, and Liwei Chan. "Slice of Light: Transparent and Integrative Transition Among Realities in a Multi-HMD-User Environment." In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST '20. Virtual Event, USA: ACM, 2020, pp. 805–817. DOI: 10.1145/3379337.3415868.

[Wan+20b] Chiu-Hsuan Wang, Seraphina Yong, Hsin-Yu Chen, Yuan-Syun Ye, and Liwei Chan. "HMD Light: Sharing In-VR Experience via Head-Mounted Projector for Asymmetric Interaction." In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST '20. Virtual Event, USA: ACM, 2020, pp. 472–486. DOI: 10.1145/3379337.3415847.

[Wei+11] Thibaut Weise, Sofien Bouaziz, Hao Li, and Mark Pauly. "Realtime Performance-Based Facial Animation." In: *ACM Transactions on Graphics (siggraph'11)* 30.4 (2011), 77:1–10. DOI: 10.1145/2010324.1964972.

[WBF20] Tim Weissker, Pauline Bimberg, and Bernd Froehlich. "Getting There Together: Group Navigation in Distributed Virtual Environments." In: *IEEE Transactions on Visualization and Computer Graphics* 26.5 (2020), pp. 1860–1870. DOI: 10.1109/TVCG.2020.2973474.

[WF21] Tim Weissker and Bernd Froehlich. "Group Navigation for Guided Tours in Distributed Virtual Environments." In: *IEEE Transactions on Visualization and Computer Graphics* 27.5 (2021), pp. 2524–2534. DOI: 10.1109/TVCG.2021.3067756.

[WKF19] Tim Weissker, Alexander Kulik, and Bernd Froehlich. "Multi-Ray Jumping: Comprehensible Group Navigation for Collocated Users in Immersive Virtual Reality." In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2019, pp. 136–144. DOI: 10.1109/VR.2019.8797807.

[Wei+18] Tim Weissker, André Kunert, Bernd Fröhlich, and Alexander Kulik. "Spatial Updating and Simulator Sickness During Steering and Jumping in Immersive Virtual Environments." In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2018, pp. 97–104. DOI: 10.1109/VR.2018.8446620.

[Wil+10] Malte Willert, Stephan Ohl, Anke Lehmann, and Oliver Staadt. "The Extended Window Metaphor for Large High-resolution Displays." In: *Proceedings of the Eurographics Conference on Virtual Environments & Second Joint Virtual Reality*. EGVE - JVRC'10. Stuttgart, Germany: Eurographics Association, 2010, pp. 69–76. DOI: 10.2312/EGVE/JVRC10/069-076.

[WG10] Nelson Wong and Carl Gutwin. "Where Are You Pointing?: The Accuracy of Deictic Pointing in CVEs." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI'10*. Atlanta, Georgia, USA: ACM, 2010, pp. 1029–1038. DOI: 10.1145/1753326.1753480.

[Xia+18] Haijun Xia, Sebastian Herscher, Ken Perlin, and Daniel Wigdor. "Spacetime: Enabling Fluid Individual and Collaborative Editing in Virtual Reality." In: *Proceedings of the Symposium on User Interface Software and Technology*. UIST '18. Berlin, Germany: ACM, 2018, pp. 853–866. DOI: 10.1145/3242587.3242597.

[XZ15] Junjie Xue and Gang Zhao. "Interactive Rendering and Modification of Massive Aircraft CAD Models in Immersive Environment." In: *Computer-Aided Design and Applications* 12.4 (2015), pp. 393–402. DOI: 10.1080/16864360.2014.997635.

[Yam+02] Toshio Yamada, Daisuke Tsubouchi, Tetsuro Ogi, and Michitaka Hirose. "Desk-sized immersive workplace using force feedback grid interface." In: *Proceedings IEEE Virtual Reality (VR)*. 2002, pp. 135–142. DOI: 10.1109/VR.2002.996516.

[Yan+22] Longqi Yang et al. "The effects of remote work on collaboration among information workers." In: *Nature Human Behaviour* 6 (2022), pp. 43–54. DOI: 10.1038/s41562-021-01196-4.

[Yao+18] Nancy Yao, Jeff Brewer, Sarah D'Angelo, Mike Horn, and Darren Gergle. "Visualizing Gaze Information from Multiple Students to Support Remote Instruction." In: *Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems*. CHI EA '18. Montreal QC, Canada: ACM, 2018, pp. 1–6. DOI: 10.1145/3170427.3188453.

[Yoo+19] Boram Yoon, Hyung-il Kim, Gun A. Lee, Mark Billinghurst, and Woontack Woo. "The Effect of Avatar Appearance on Social Presence in an Augmented Reality Remote Collaboration." In: *Proceedings of IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2019, pp. 547–556. DOI: 10.1109/VR.2019.8797719.

[Yu+22] Kevin Yu, Ulrich Eck, Frieder Pankratz, Marc Lazarovici, Dirk Wilhelm, and Nassir Navab. "Duplicated Reality for Co-located Augmented Reality Collaboration." In: *IEEE Transactions on Visualization and Computer Graphics* 28.5 (2022), pp. 2190–2200. DOI: 10.1109/TVCG.2022.3150520.

[Yu+10] Lingyun Yu, Pjotr Svetachov, Petra Isenberg, Maarten H. Everts, and Tobias Isenberg. "FI3D: Direct-Touch Interaction for the Exploration of 3D Scientific Visualization Spaces." In: *IEEE Transactions on Visualization and Computer Graphics* 16.6 (2010), pp. 1613–1622. DOI: 10.1109/TVCG.2010.157.

[Zha+23] Lei Zhang, Ashutosh Agrawal, Steve Oney, and Anhong Guo. "VRGit: A Version Control System for Collaborative Content Creation in Virtual Reality." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '23. Hamburg, Germany: ACM, 2023. DOI: 10.1145/3544548.3581136.

[Zha+04] Li Zhang, Noah Snavely, Brian Curless, and Steven M. Seitz. "Spacetime Faces: High Resolution Capture for Modeling and Animation." In: *ACM Transactions on Graphics (siggraph'04)* 23.3 (2004), pp. 548–558. DOI: 10.1145/1015706.1015759.

[Zha+21] Yiran Zhang, Sy-Thanh Ho, Nicolas Ladévèze, Huyen Nguyen, Cédric Fleury, and Patrick Bourdot. "In Touch with Everyday Objects: Teleportation Techniques in Virtual Environments Supporting Tangibility." In: *Workshop on Everyday Virtual Reality (WEVR) at the IEEE Conference on Virtual Reality and 3D User Interfaces*. Virtual Event, 2021, pp. 278–283. DOI: `10.1109/VRW52623.2021.00057`.

[Zha+19] Yiran Zhang, Nicolas Ladévèze, Cédric Fleury, and Patrick Bourdot. "Switch Techniques to Recover Spatial Consistency Between Virtual and Real World for Navigation with Teleportation." In: *Proceedings of the EuroVR International Conference*. Vol. 11883. Lecture Notes in Computer Science. Tallinn, Estonia: Springer, 2019, pp. 3–23. DOI: `10.1007/978-3-030-31908-3_1`.

[Zha+20] Yiran Zhang, Nicolas Ladévèze, Huyen Nguyen, Cédric Fleury, and Patrick Bourdot. "Virtual Navigation Considering User Workspace: Automatic and Manual Positioning before Teleportation." In: *Proceedings of the Symposium on Virtual Reality Software and Technology*. VRST'20. Virtual Event: ACM, 2020. DOI: `10.1145/3385956.3418949`.

[Zha+22] Yiran Zhang, Huyen Nguyen, Nicolas Ladevèze, Cédric Fleury, and Patrick Bourdot. "Virtual Workspace Positioning Techniques during Teleportation for Co-located Collaboration in Virtual Reality using HMDs." In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces*. IEEE VR'22. 2022, pp. 674–682. DOI: `10.1109/VR51125.2022.00088`.

CURRICULUM VITAE

Cédric FLEURY - *https://cedricfleury.fr*

Current Situation & Research Positions

2021 - present	**Assistant Professor**, IMT Atlantique / Lab-STICC, Brest (France) *Member of the **INUIT** research group*
2013 - 2021	**Assistant Professor**, Université Paris-Saclay / LISN, Orsay (France) *Member of the **ex)situ** Inria research group*
2012 - 2013	**Postdoctoral researcher**, University of North Carolina (UNC), Chapel Hill (USA) Visiting researcher at Nanyang Technological University (NTU), Singapour *Supervisors: Henry Fuchs, Tat Jen Cham*
2008 - 2012	**Ph.D. student**, IRISA / Inria Rennes (France), funded by the "ANR Collaviz" project *Supervisors: Bruno Arnaldi (thesis director), Thierry Duval, Valérie Gouranton*
2011 Sept. - Dec.	**Visiting Ph.D. student**, University College of London (UK) - *supervised by Anthony Steed*
2008 Feb. - June	**Master student**, IRISA / Inria Rennes (France) - *supervised by Thierry Duval*
2007 Feb. - June	**Master student**, IRISA / Inria Rennes (France) - *supervised by Thierry Duval*
2006 July - Sept.	**Intern**, INSA de Rennes (France) - *supervised by Pierre-Yves Glorennec*

Education

2012	**Ph.D. in Computer Science**, INSA de Rennes (France)
2008	**Master of Research degree** in *Interaction, Image Processing and AI*, INSA de Rennes (France)
2008	**Master of Science degree** in *Computer Science*, INSA de Rennes (France)
2004	**Bachelor of Science degree in Mathematics**, Université de Bretagne Sud, Lorient (France)
2002	**Baccalauréat in sciences** (high school diploma), Lycée St Charles, St Brieuc, (France)

Teaching

Courses at undergraduate level: intro. to computer science, object-oriented programming, intro. to HCI
Courses at graduate level: VR and AR, Advanced HCI, software engineering for HCI, CSCW

2021 - present	**Assistant Professor of Computer Science** at IMT Atlantique engineering school and in the human-computer interaction Master at Université Bretagne Occidentale
2021 - present	**Co-chair of the HCI Major** (Master's level), IMT Atlantique
2013 - 2021	**Assistant Professor of Computer Science** at Polytech Paris-Saclay engineering school and in the human-computer interaction Masters at Université Paris-Saclay
2013 - 2021	**Head of a technological platform** used for teaching mixte reality, Polytech Paris-Saclay
2019 - 2021	**Chair of the 5th year** of engineering school in computer science, Polytech Paris-Saclay
2017 - 2019	**Chair of the 3rd year** of apprenticeship in computer science, Polytech Paris-Saclay
2013 - 2018	**Head of internships** in the human-computer interaction Masters, Université Paris-Saclay
2006 - 2011	Teaching assistant, INSA de Rennes (France)
2008 - 2010	Teaching assistant, Université de Rennes 1 (France)

Research

My research interests are in the fields of human-computer interaction, computer-supported cooperative work, and mixte reality including virtual and augmented reality. I am interested in the collaboration and interaction of multiple users in large interactive spaces such as wall-sized displays, immersive virtual reality systems and augmented reality setups. I worked on various projects on video-conferencing, telepresence, collaborative virtual environments and 3D interaction.

Supervision: I supervised 14 master’s level interns, 6 engineers and 6 Ph.D. students (2 Ph.D.s are in progress)

Ph.D. students (6)

2022 - present	Aurélien Léchappé, supervised at 40% with C. Dumas and M. Chollet *"Modeling common ground knowledge for real-time analysis of collaboration in a virtual env."*
2020 - present	Thomas Rinnert, supervised at 30% with T. Duval et B. Thomas *"Perceiving distant collaborative activity with mixed reality"*
2019 - 2023	Arthur Fages, supervised at 50% with T. Tsandilas - **thesis defended** *"Supporting collaborative 3D modeling through augmented-reality spaces"*
2017 - 2021	Yiran Zhang, supervised at 50% with P. Bourdot - **thesis defended** *"Telepresence for remote and heterogeneous collaborative virtual environments"*
2015 - 2019	Yujiro Okuya, supervised at 50% with P. Bourdot - **thesis defended** "*CAD modification techniques for design reviews on heterogeneous interactive systems"*
2014 - 2017	Ignacio Avellino, supervised at 80% with M. Beaudouin-Lafon - **thesis defended** *"Supporting collaborative practices across wall-sized displays with video-mediated communication"*

Master’s level interns (14)

- Clément Jézéquel (M2), co-supervised with E. Peillard (2023). *Keywords: remote collab., AR*
- Michele Romani (M1), co-supervised with M. Beaudouin-Lafon (2019). *Keywords: telepresence, AR*
- Clément Sauvard (M2), co-supervised with T. Tsandilas (2019). *Keywords: 2D/3D interaction, AR*
- Cyril Crebouw (M2), co-supervised with M. Beaudouin-Lafon (2019). *Keywords: telepresence, WSD*
- Antonin Cheymol (M1), supervised at 100% (2019). *Keywords: 3D interaction, VR*
- Jiannan Li (visiting PhD), co-supervised with M. Beaudouin-Lafon (2018). *Keywords: telepresence, WSD*
- Kévin Ahson (M1), co-supervised with T. Tsandilas (2018). *Keywords: collaborative interaction, AR*
- Krishnan Chandran (M2), co-supervised with T. Tsandilas (2018). *Keywords: collaborative interaction, AR*
- Brennan Jones (M2), co-supervised with I. Avellino et M. Beaudouin-Lafon (2016). *Keywords: telepresence, WSD*
- Jean-Baptiste Louvet (M2), supervised at 100% (2015). *Keywords: 3D interaction, WSD*
- Ignacio Avellino (M2), co-supervised with M. Beaudouin-Lafon (2014). *Keywords: telepresence, WSD*
- Hugo Marchadour (M1), co-supervised with T. Duval et V. Gouranton (2011). *Keywords: visualization, VR*
- Florent Goetz (M1), co-supervised with T. Duval et V. Gouranton (2011). *Keywords: collaborative interaction, VR*
- Charles Perin (M1), co-supervised with T. Duval (2010). *Keywords: 3D interaction, VR*

Engineers (6)

- Léo Colombaro (intern), co-supervised with O. Gladin (2018)
- Gabriel Tézier (2014), Amani Kooli, Jonathan Thorpe, Rémi Hellequin (2014 - 2016) and Lawrence Fyfe (2016 - 2018), co-supervised with J. Vézien, O. Gladin, S. Huot and M. Beaudouin-Lafon in the context of DIGISCOPE

Research projects

2022 - present	Industrial chair "Région Pays-de-la-Loire" - Ph.D. funding of Aurélien Léchappé
2021 - present	Participation in the EquipEx+ "Continuum" (scientific and technical committees)
2014 - 2020	Co-chair of the **technical committee** (with J. Vézien) and budget monitoring of the EquipEx "DIGISCOPE" partially funded by the French National Research Agency (ANR) (22M€)
2015 - 2019	Co-investigator of "SensoMotorCVE" project (with P. Bourdot) - Ph.D. funding of Y. Okuya by Labex Digicosme - ANR (109K€)
2018	Participation in "VR-BatIM" project (main investigator: J. Vézien) - Univ. Paris-Saclay (20K€) *"Virtual reality for the interaction with building information model at Paris-Saclay"*
2017	Participation in "ARSCPMD" project (main investigator: T. Tsandilas) - Univ. Paris-Saclay (8K€) *"An augmented-reality system for collaborative physical modeling and design"*
2015	Main investigator of an "Attractivité" project - Université Paris-Saclay (7,3K€) *"Telepresence systems preserving eye contact between remote users located in large interactive spaces"*

2012 - 2013	Participation in the *BeingThere Center,* a collaborative project between University of North Carolina (UNC) at Chapel Hill (USA), Nanyang Technological University (NTU) in Singapour and Swiss Federal Institute of Technology (ETH) in Zurich (Switzerland)
2009 - 2012	Ph.D. student for the ANR "Collaviz" project
2007 - 2010	Intern, then Ph.D. student for the RNTL / ANR "Part@ge" project

Organization and research evaluation

- Co-chair of the doctoral consortium at IEEE ISMAR 2022 conference
- Co-chair of the program committees for the demonstrations at IHM 2021 conference
- Co-chair of the program committees for the "work in progress" at IHM 2018 conference
- Member of the program committees of the following conferences: ACM VRST 2019, EuroVR (2017 and 2020), GI 2016, IEEE VR 2016, 3DCVE@IEEE VR (2015 - 2018), GRAPP (2014 - 2016)
- Reviewer for the following journals: TVCG, Frontiers, CG&A, JOCCH, JVRB
- Reviewer for the following conferences: ACM CHI, ACM UIST, ACM VRST, ACM SIGGRAPH ASIA, IEEE VR, IEEE ISMAR, 3DUI, 3DVIS@IEEE VIS, 3DCVE@IEEE VR, CGI, JVRC, IHM
- Reviewer for project proposals of the French National Research Agency (ANR) (2014, 2015, 2020, 2023)
- Member of the PhD. committee of Romain Terrier (IRISA - Inria Rennes) (2020)

Professional service

- Elected member of the **LRI lab council** (joint research lab between Univ. Paris-Saclay & CNRS) (2014 - 2020)
- Member of a **hiring committee for an assistant professor position** at Université Paris-Saclay (2019)

Other experiences

2002 - 2009	**High level athlete in sailing** - registered on the list of the French Ministry of Sports International level in Olympic sailing (49er) and offshore racing ("Tour de France" sailing races)
2002 - 2005	**Treasurer of the association 29er - 49er France:** this association federates French competitors sailing on 49er and 29er boats, and organizes international events.

PUBLICATIONS

BOOK CHAPTER

[Oku+21] Yujiro Okuya, Nicolas Ladévèze, Olivier Gladin, **Cédric Fleury**, and Patrick Bourdot. “Collaborative VR-CAD for Industrial Product Design: CAD Parameter Modification with 3D Interaction on Heterogeneous Immersive Platforms.” In: *Manufacturing in the Era of 4th Industrial Revolution*. Ed. by Monica Bordegoni, Satyandra K. Gupta, and James Ritchie. Vol. 3. World Scientific, 2021. Chap. 2, pp. 17–47. DOI: `10.1142/9789811222863_0002`.

JOURNALS

[FFT22b] Arthur Fages, **Cédric Fleury**, and Theophanis Tsandilas. “Understanding Multi-View Collaboration between Augmented Reality and Remote Desktop Users.” In: *Proceedings of the ACM on Human-Computer Interaction* 6. CSCW2 (2022). DOI: `10.1145/3555607`.

[Oku+18a] Yujiro Okuya, Nicolas Ladévèze, **Cédric Fleury**, and Patrick Bourdot. “ShapeGuide: Shape-Based 3D Interaction for Parameter Modification of Native CAD Data.” In: *Frontiers in Robotics and AI* 5 (2018). DOI: `10.3389/frobt.2018.00118`.

[Duv+14] Thierry Duval, Huyen Nguyen, **Cédric Fleury**, Alain Chauffaut, Georges Dumont, and Valérie Gouranton. “Improving awareness for 3D virtual collaboration by embedding the features of users’ physical environments and by augmenting interaction tools with cognitive feedback cues.” In: *Journal on Multimodal User Interfaces* 8 (2014), pp. 187–197. DOI: `10.1007/s12193-013-0134-z`.

INTERNATIONAL CONFERENCES

[Zha+22] Yiran Zhang, Huyen Nguyen, Nicolas Ladevèze, **Cédric Fleury**, and Patrick Bourdot. “Virtual Workspace Positioning Techniques during Teleportation for Co-located Collaboration in Virtual Reality using HMDs.” In: *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces*. IEEE VR’22. 2022, pp. 674–682. DOI: `10.1109/VR51125.2022.00088`.

[Oku+20] Yujiro Okuya, Olivier Gladin, Nicolas Ladévèze, **Cédric Fleury**, and Patrick Bourdot. “Investigating Collaborative Exploration of Design Alternatives on a Wall-Sized Display.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI’20. Virtual Event: ACM, 2020, pp. 1–12. DOI: `10.1145/3313831.3376736`.

[Zha+20] Yiran Zhang, Nicolas Ladévèze, Huyen Nguyen, **Cédric Fleury**, and Patrick Bourdot. "Virtual Navigation Considering User Workspace: Automatic and Manual Positioning before Teleportation." In: *Proceedings of the Symposium on Virtual Reality Software and Technology*. VRST'20. Virtual Event: ACM, 2020. DOI: `10.1145/3385956.3418949`.

[Le+19] Khanh-Duy Le, Ignacio Avellino, **Cédric Fleury**, Morten Fjeld, and Andreas M. Kunz. "GazeLens: Guiding Attention to Improve Gaze Interpretation in Hub-Satellite Collaboration." In: *Proceedings of the IFIP TC13 International Conference On Human-Computer Interaction (INTERACT)*. Vol. 11747. Lecture Notes in Computer Science. Paphos, Cyprus: Springer, 2019, pp. 282–303. DOI: `10.1007/978-3-030-29384-0\_18`.

[Zha+19] Yiran Zhang, Nicolas Ladévèze, **Cédric Fleury**, and Patrick Bourdot. "Switch Techniques to Recover Spatial Consistency Between Virtual and Real World for Navigation with Teleportation." In: *Proceedings of the EuroVR International Conference*. Vol. 11883. Lecture Notes in Computer Science. Tallinn, Estonia: Springer, 2019, pp. 3–23. DOI: `10.1007/978-3-030-31908-3_1`.

[Fyf+18] Lawrence Fyfe, Olivier Gladin, **Cédric Fleury**, and Michel Beaudouin-Lafon. "Combining Web Audio Streaming, Motion Capture, and Binaural Audio in a Telepresence System." In: *Proceedings of the International Web Audio conference*. WAC'18. Berlin, Germany, 2018. URL: `https://hal.science/hal-01957843`.

[Ave+17] Ignacio Avellino, **Cédric Fleury**, Wendy E. Mackay, and Michel Beaudouin-Lafon. "CamRay: Camera Arrays Support Remote Collaboration on Wall-Sized Displays." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI'17. Denver, Colorado, USA: ACM, 2017, pp. 6718–6729. DOI: `10.1145/3025453.3025604`.

[LF16] Jean-Baptiste Louvet and **Cédric Fleury**. "Combining Bimanual Interaction and Teleportation for 3D Manipulation on Multi-Touch Wall-Sized Displays." In: *Proceedings of the Symposium on Virtual Reality Software and Technology*. VRST'16. Munich, Germany: ACM, 2016, pp. 283–292. DOI: `10.1145/2993369.2993390`.

[AFB15] Ignacio Avellino, **Cédric Fleury**, and Michel Beaudouin-Lafon. "Accuracy of Deictic Gestures to Support Telepresence on Wall-Sized Displays." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI'15. Seoul, Republic of Korea: ACM, 2015, pp. 2393–2396. DOI: `10.1145/2702123.2702448`.

[Fle+14] **Cédric Fleury**, Tiberiu Popa, Tat Jen Cham, and Henry Fuchs. "Merging Live and pre-Captured Data to support Full 3D Head Reconstruction for Telepresence." In: *Proceedings of the Conference of the European Association for Computer Graphics*. EG'14. Strasbourg, France, 2014, pp. 9–12. DOI: `10.2312/egsh.20141002`.

[NDF13b] Thi Thuong Huyen Nguyen, Thierry Duval, and **Cédric Fleury**. "Guiding Techniques for Collaborative Exploration in Multi-Scale Shared Virtual Environments." In: *Proceedings of the International Conference* *on Computer Graphics Theory and Applications*. GRAPP'13. Barcelone, Spain, 2013, pp. 327–336. URL: https://hal.science/hal-00755313.

[Duv+12] Thierry Duval, Huyen Nguyen, **Cédric Fleury**, Alain Chauffaut, Georges Dumont, and Valérie Gouranton. "Embedding the features of the users' physical environments to improve the feeling of presence in collaborative Virtual Environments." In: *Proceedings of the IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. 2012, pp. 243–248. DOI: 10.1109/CogInfoCom.2012.6421987.

[Fle+12] **Cédric Fleury**, Thierry Duval, Valérie Gouranton, and Anthony Steed. "Evaluation of Remote Collaborative Manipulation for Scientific Data Analysis." In: *Proceedings of the Symposium on Virtual Reality Software and Technology*. VRST'12. Toronto, Ontario, Canada: ACM, 2012, pp. 129–136. DOI: 10.1145/2407336.2407361.

[DF11b] Thierry Duval and **Cédric Fleury**. "PAC-C3D: A New Software Architectural Model for Designing 3D Collaborative Virtual Environments." In: *Proceedings of the International Conference on Artificial Reality and Telexistence*. ICAT'11. Osaka, Japan, 2011, pp. 53–60. URL: https://hal.science/inria-00636143.

[Fle+10a] **Cédric Fleury**, Alain Chauffaut, Thierry Duval, Valérie Gouranton, and Bruno Arnaldi. "A Generic Model for Embedding Users' Physical Workspaces into Multi-Scale Collaborative Virtual Environments." In: *Proceedings of the International Conference on Artificial Reality and Telexistence*. ICAT'10. Adelaide, Australia, 2010. URL: https://hal.science/inria-00534096.

[Fle+10b] **Cédric Fleury**, Thierry Duval, Valérie Gouranton, and Bruno Arnaldi. "A New Adaptive Data Distribution Model for Consistency Maintenance in Collaborative Virtual Environments." In: *Proceedings of the Joint Virtual Reality Conference of EuroVR - EGVE - VEC*. EGVE - JVRC'10. Stuttgart, Germany: Eurographics Association, 2010, pp. 29–36. DOI: 10.2312/EGVE/JVRC10/029-036.

[DF09] Thierry Duval and **Cédric Fleury**. "An Asymmetric 2D Pointer/3D Ray for 3D Interaction within Collaborative Virtual Environments." In: *Proceedings of the International Conference on 3D Web Technology*. Web3D'09. Darmstadt, Germany: ACM, 2009, pp. 33–41. DOI: 10.1145/1559764.1559769.

WORKSHOP AND LATE-BREAKING WORK PAPERS

[Léc+23] Aurélien Léchappé, Aurélien Milliat, **Cédric Fleury**, Mathieu Chollet, and Cédric Dumas. "Characterization of Collaboration in a Virtual Environment with Gaze and Speech Signals." In: *Companion Publication of the International Conference on Multimodal Interaction*. ICMI '23 Companion. Paris, France: ACM, 2023, pp. 21–25. DOI: 10.1145/3610661.3617149.

[Zha+21] Yiran Zhang, Sy-Thanh Ho, Nicolas Ladévèze, Huyen Nguyen, **Cédric Fleury**, and Patrick Bourdot. “In Touch with Everyday Objects: Teleportation Techniques in Virtual Environments Supporting Tangibility.” In: *Workshop on Everyday Virtual Reality (WEVR) at the IEEE Conference on Virtual Reality and 3D User Interfaces*. Virtual Event, 2021, pp. 278–283. DOI: 10.1109/VRW52623.2021.00057.

[Oku+18b] Yujiro Okuya, Nicolas Ladévèze, Olivier Gladin, **Cédric Fleury**, and Patrick Bourdot. “Distributed Architecture for Remote Collaborative Modification of Parametric CAD Data.” In: *Workshop on Collaborative Virtual Environments (3DCVE) at the IEEE Conference on Virtual Reality*. Reutlingen, Germany, 2018, pp. 1–4. DOI: 10.1109/3DCVE.2018.8637112.

[Fle+15a] **Cédric Fleury**, Ignacio Avellino, Michel Beaudouin-Lafon, and Wendy E. Mackay. “Telepresence Systems for Large Interactive Spaces.” In: *Workshop on Everyday Telepresence: Emerging Practices and Future Research Directions at the ACM conference on Human Factors in Computing Systems*. Seoul, South Korea, 2015. URL: https://hal.science/hal-01242304.

[Fle+15b] **Cédric Fleury**, Nicolas Férey, Jean-Marc Vézien, and Patrick Bourdot. “Remote collaboration across heterogeneous large interactive spaces.” In: *Workshop on Collaborative Virtual Environments (3DCVE) at the IEEE Conference on Virtual Reality*. 2015, pp. 9–10. DOI: 10.1109/3DCVE.2015.7153591.

[Fle+10c] **Cédric Fleury**, Thierry Duval, Valérie Gouranton, and Bruno Arnaldi. “Architectures and Mechanisms to Maintain efficiently Consistency in Collaborative Virtual Environments.” In: *Workshop on Software Engineering and Architectures for Realtime Interactive Systems (SEARIS) at the IEEE Conference on Virtual Reality*. Waltham, United States, 2010. URL: https://hal.science/inria-00534082.

POSTERS AND DEMONSTRATIONS

[FFT22a] Arthur Fages, **Cédric Fleury**, and Theophanis Tsandilas. “ARgus : système multi-vues pour collaborer à distance avec un utilisateur en réalité augmentée.” In: *Proceedings of the conférence francophone sur l’Interaction Homme-Machine*. Demonstration. Namur, Belgium, 2022. URL: https://hal.science/hal-03762816.

[Oku+17] Yujiro Okuya, Nicolas Ladévèze, **Cédric Fleury**, and Patrick Bourdot. “VR-CAD Framework for Parametric Data Modification with a 3D Shape-based Interaction.” In: *Proceedings of the EuroVR International Conference (poster)*. Laval, France, 2017. URL: https://hal.science/hal-04327386.

[NDF13a] Thi Thuong Huyen Nguyen, Thierry Duval, and **Cédric Fleury**. “Demonstration of Guiding Techniques for Collaborative Exploration in Multi-Scale Shared Virtual Environments.” In: *Proceedings of the conférence francophone sur l’Interaction Homme-Machine (demo)*. Demonstration. Bordeaux, France, 2013. URL: https://inria.hal.science/hal-00879475.

[NFD12] Thi Thuong Huyen Nguyen, **Cédric Fleury**, and Thierry Duval. "Collaborative exploration in a multi-scale shared virtual environment." In: *Proceedings of the Symposium on 3D User Interfaces*. Demonstration. 2012, pp. 181–182. DOI: 10.1109/3DUI.2012.6184221.

[DF11a] Thierry Duval and **Cédric Fleury**. "Collaboration between Networked Heterogeneous 3D Viewers through a PAC-C3D Modeling of the Shared Virtual Environment." In: *Proceedings of the International Conference on Artificial Reality and Telexistence*. Demonstration. Osaka, Japan, 2011. URL: https://hal.science/inria-00638638v1.

[Dup+10] Florent Dupont, Thierry Duval, **Cédric Fleury**, Julien Forest, Valérie Gouranton, Pierre Lando, Thibaut Laurent, Guillaume Lavoué, and Alban Schmutz. "Collaborative Scientific Visualization: The COLLAVIZ Framework." In: *Proceedings of the Joint Virtual Reality Conference of EuroVR - EGVE - VEC*. Demonstration. Fellbach, Germany, 2010. URL: https://hal.science/inria-00534105.

[Duv+08] Thierry Duval, **Cédric Fleury**, Bernard Nouailhas, and Laurent Aguerreche. "Collaborative Exploration of 3D Scientific Data." In: *Proceedings of the Symposium on Virtual Reality Software and Technology (VRST'08)*. Demonstration. Bordeaux, France: ACM, 2008, pp. 303–304. DOI: 10.1145/1450579.1450664.

NATIONAL CONFERENCES

[FFT23] Arthur Fages, **Cédric Fleury**, and Theophanis Tsandilas. "Conception collaborative au travers de versions parallèles en réalité augmentée." In: *Proceedings of the Conference on l'Interaction Humain-Machine*. IHM '23. TROYES, France: ACM, 2023. DOI: 10.1145/3583961.3583978.

[FDC08] **Cédric Fleury**, Thierry Duval, and Alain Chauffaut. "Cabine Virtuelle d'Immersion (CVI) : naviguer et interagir en immersion dans les univers virtuels collaboratifs multi-échelles." In: *Journées de l'Association Française de Réalité Virtuelle (AFRV)*. Bordeaux, France, 2008. URL: https://hal.science/inria-00433843.

SELECTED PUBLICATIONS

p. 117 J.-B. Louvet, C. Fleury (2016). Combining Bimanual Interaction and Teleportation for 3D Manipulation on Multi-Touch Wall-sized Displays. Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST), 8 pages.

p. 127 Y. Okuya, O. Gladin, N. Ladévèze, C. Fleury, P. Bourdot (2020). Investigating Collaborative Exploration of Design Alternatives on a Wall-Sized Display. Proceedings of the ACM conference on Human Factors in Computing Systems (CHI), 12 pages.

p. 139 Y. Zhang, N. Ladévèze, H. Nguyen, C. Fleury, P. Bourdot (2020). Virtual Navigation considering User Workspace: Automatic and Manual Positioning before Teleportation. Proc. of the ACM Symposium on Virtual Reality Software and Technology (VRST), 10 pages.

p. 149 Y. Zhang, T.T.H. Nguyen, N. Ladévèze, C. Fleury, P. Bourdot (2022). Virtual Workspace Positioning Techniques during Teleportation for Co-located Collaboration in Virtual Reality using HMDs, Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces (IEEE VR), 9 pages.

p. 158 I. Avellino, C. Fleury, M. Beaudouin-Lafon (2015). Accuracy of deictic gestures to support telepresence on wall-sized displays. Proceedings of the ACM conference on Human Factors in Computing Systems (CHI), 4 pages.

p. 162 I. Avellino, C. Fleury, W. Mackay, M. Beaudouin-Lafon (2017). CamRay: Camera Arrays Support Remote Collaboration on Wall- Sized Displays. Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI), 12 pages.

p. 174 K.-D. Le, I. Avellino, C. Fleury, M. Fjeld, A. Kunz (2019). GazeLens: Guiding Attention to Improve Gaze Interpretation in Hub-Satellite Collaboration. Proc. of the IFIP TC13 Conference on Human-Computer Interaction (INTERACT), 21 pages.

p. 195 A. Fages, C. Fleury, T. Tsandilas (2022). Understanding Multi-View Collaboration between Augmented Reality and Remote Desktop Users, Proc. of the ACM Conference on Computer-Supported Cooperative Work (CSCW), 27 pages.

Titre : Favoriser la collaboration dans les grands espaces interactifs

Mots clés : interaction humain-machine, travail coopératif assisté par ordinateur, réalité virtuelle, réalité augmentée, téléprésence, interaction 3D

Résumé : Avec la croissance exponentielle de la quantité et de la complexité des données numériques produites par notre société, le besoin d'outils informatiques pour collaborer n'a jamais été aussi important. Permettre à des groupes d'utilisateurs de manipuler, d'analyser et de comprendre ces données, tout en conservant le contrôle sur la façon dont l'intelligence artificielle les traite, est devenu un défi majeur. Dans ce contexte, mes recherches étudient comment les grands espaces interactifs, tels que les murs d'images, les systèmes immersifs de réalité virtuelle ou les espaces de réalité augmentée, peuvent favoriser la collaboration entre les utilisateurs.

La première partie de mon travail explore de nouveaux paradigmes d'interaction permettant aux utilisateurs de maîtriser les caractéristiques inhabituelles des grands espaces interactifs. Au-delà de l'interaction à un niveau individuel, il s'agit d'étudier comment ces systèmes peuvent favoriser la collaboration entre utilisateurs co-localisés. La seconde partie de mon travail porte sur la collaboration à distance entre espaces interactifs. Elle propose à la fois des solutions techniques pour connecter des plateformes hétérogènes, et des solutions pour favoriser la perception mutuelle et la communication entre les collaborateurs distants. Plutôt que de chercher à reproduire la collaboration dans le monde physique, mon travail propose d'aller au-delà en exploitant les capacités numériques et le grand espace physique qui entoure les utilisateurs.

Mes travaux futurs se concentreront sur comment exploiter au mieux le continuum de la réalité mixte pour permettre à des utilisateurs d’interagir et de collaborer à différents niveaux de ce continuum. L'objectif principal est de pouvoir s'adapter aux différentes phases de la collaboration dans des situations hybrides impliquant à la fois des participants co-localisés et distants.

Title: Supporting Collaboration in Large Interactive Spaces

Keywords: human-computer interaction, computer-supported cooperative work, virtual reality, augmented reality, telepresence, 3D interaction

Abstract: As the quantity and complexity of digital data produced by our society grow exponentially, the need for computer-supported collaboration has never been higher. Empowering groups of users to manipulate, analyze and understand this data, while preserving control over how artificial intelligence processes it, has become a major challenge. In this context, my research investigates how large interactive spaces, such as wall-sized displays, immersive virtual reality systems or augmented reality spaces, can foster collaboration among users.

A first part of my work investigates new interaction paradigms that provide users with the ability to master the unusual characteristics of such large interactive spaces. Beyond individual interaction, it investigates how these systems can foster co-located collaboration by providing appropriate collaborative interaction among users. A second part of my work focuses on remote collaboration across large interactive spaces. It explores technical solutions to connect heterogeneous platforms, as well as telepresence systems providing appropriate awareness and communication cues among the remote collaborators. Rather than mimicking collaboration in the physical world, it aims to push collaboration beyond “being there” by leveraging digital cues and taking advantage of the large physical space surrounding users.

My future research will concentrate on exploiting the mixed reality continuum to enable collaborators to interact across time and space by seamlessly transitioning between heterogeneous interaction modalities. The overall objective is to support the different phases of a collaboration in hybrid situations, involving both co-located and remote participants.