



HAL
open science

Collaborative Design in Augmented Reality: Supporting multiple views and parallel versions of the design space

Arthur Fages

► To cite this version:

Arthur Fages. Collaborative Design in Augmented Reality: Supporting multiple views and parallel versions of the design space. Human-Computer Interaction [cs.HC]. Université Paris-Saclay, 2023. English. NNT: 2023UPASG027 . tel-04098984

HAL Id: tel-04098984

<https://theses.hal.science/tel-04098984>

Submitted on 16 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Collaborative Design in Augmented Reality : Supporting Multiple Views and Parallel Versions of the Design Space

*Conception collaborative en Réalité Augmentée :
prise en charge de vues multiples
et de versions parallèles de l'espace de conception*

Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 580,
sciences et technologies de l'information et de la communication (STIC)
Spécialité de doctorat : Informatique
Graduate School : Informatique et sciences du numérique.
Réfèrent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche **Laboratoire interdisciplinaire des sciences
du numérique (Université Paris-Saclay, CNRS)**
sous la direction de **Cédric FLEURY**, Maître de Conférence
et la co-direction de **Theophanis TSANDILAS**, Chargé de Recherche

Thèse soutenue à Paris-Saclay, le 11 Avril 2023, par

Arthur FAGES

Composition du jury

Membres du jury avec voix délibérative

Anastasia BEZERIANOS Professeur des universités, Université Paris-Saclay	Présidente
Ferran ARGELAGUET Chargé de recherche, INRIA	Rapporteur & Examineur
Géry CASIEZ Professeur des universités, Université de Lille	Rapporteur & Examineur
Morten FJELD Professeur, Chalmers University of Technology & University of Bergen	Examineur

Titre : Conception collaborative en Réalité Augmentée : prise en charge de vues multiples et de versions parallèles de l'espace de conception

Mots clés : Interaction Humain-Machine, Collaboration, Réalité Augmentée

Résumé : La collaboration est un aspect essentiel du processus de conception. Au cours de ses premières phases, plusieurs concepteurs se réunissent pour partager leurs connaissances et leur expertise afin de résoudre les problèmes de conception. Ils génèrent, évaluent et comparent de nouvelles idées pour trouver les meilleures solutions possibles. Les technologies de réalité augmentée (RA) peuvent rendre ce processus encore plus efficace en fournissant des outils de création numérique 3D qui superposent un contenu virtuel 3D aux objets physiques.

Ma thèse explore comment les technologies de RA peuvent améliorer les tâches de conception collaborative. Bien que les systèmes de RA imposent d'ancrer le contenu virtuel dans le monde réel, je m'intéresse à élargir les limites de la conception d'interaction au-delà de ce qui est possible dans le monde physique. Pour atteindre cet objectif, j'étudie cette approche dans des situations de collaboration co-localisée et à distance.

Lorsque les concepteurs sont co-localisés dans un même environnement de RA, ils partagent le même espace physique et les mêmes objets. Toutefois, cela peut s'avérer problématique pour explorer rapidement de nombreuses idées, car ils risquent d'occuper le même espace et de se déranger mutuellement. De plus, le fait d'avoir ses créations visibles par les autres peut être une source d'inhibition.

Pour surmonter ces limites, je propose un cadre conceptuel permettant à plusieurs versions du contenu virtuel de coexister dans des espaces virtuels parallèles. Les concepteurs peuvent ainsi désynchroniser partiellement ou totalement leur environnement virtuel pour explorer indépendamment plusieurs conceptions alternatives, puis le synchroniser à nouveau pour partager et comparer leurs conceptions avec celles des autres. Pour illustrer cette approche, je présente un scénario dans lequel deux concepteurs dessinent une robe virtuelle

en 3D sur un mannequin de couture physique.

Les concepteurs sont également amenés à travailler à distance dans de nombreuses situations. Les outils de vidéoconférence peuvent faciliter ce type de collaboration et sont devenus extrêmement populaires pendant la pandémie de COVID-19 mais ils reposent principalement sur le point de vue d'une seule caméra. Ils ne donnent pas aux utilisateurs la liberté d'inspecter l'espace de travail distant car cette caméra est généralement fixe dans l'espace ou son positionnement n'est pas contrôlable. Des recherches antérieures dans le domaine de l'Interaction Humain-Machine (IHM) ont exploré l'utilisation de reconstruction 3D pour aller au-delà des outils reposant sur la vidéo. Cependant, ces techniques peuvent être complexes à mettre en place, nécessitent une acquisition de grande qualité et un réseau avec une large bande passante.

Pour explorer des alternatives à la reconstruction 3D, j'étudie des solutions qui s'appuient, de façon complémentaire, sur des représentations vidéo augmentées et virtuelles, ainsi que sur différents points de vue. J'étudie d'abord les compromis qui peuvent exister entre une vue augmentée à la première personne, une vue augmentée à la troisième personne et une vue entièrement virtuelle. Je présente ensuite ARgus, un système de communication vidéo multi-vues qui combine ces trois vues grâce à des outils interactifs de navigation, de prévisualisation, de pointage et d'annotation. Je présente ensuite une deuxième étude utilisateurs qui observe comment 12 participants ont utilisé ARgus pour communiquer des instructions à un utilisateur en RA afin de placer des meubles physiques miniatures dans une maquette virtuelle de maison. Les résultats suggèrent qu'ARgus offre une certaine flexibilité, permettant aux utilisateurs distants de vérifier plus efficacement les contraintes spatiales et réduisant leur besoin d'avoir recours à des instructions verbales.

Enfin, je discute les limites de l'approche, qui consiste à étendre les frontières de la conception d'interaction au-delà de ce qui est possible dans le monde physique. Je propose en-

suite des remarques générales sur la conception d'interaction dans les environnements de RA et suggère des opportunités pour des travaux de recherche futurs.

Title : Collaborative Design in Augmented Reality : Supporting Multiple Views and Parallel Versions of the Design Spaces

Keywords : Human Computer Interaction, Collaboration, Augmented Reality

Abstract : Collaboration is a key aspect of the design process. In its early stage, multiple designers usually share their knowledge to solve design problems by generating, evaluating, and analyzing new ideas. Using Augmented Reality (AR) technology, they benefit from 3D digital creation tools in the context of the real world, by overlaying and sharing a 3D virtual space on top of a physical space.

AR designers can be co-located by perceiving the same physical spaces and physical objects. When they want to quickly explore many ideas before discussing them with others during the early design stages, sharing the same virtual space can be problematic. Indeed, designers may compete for space when creating or evaluating more than one idea at the same time. Likewise, the visibility of their creation may be a source of inhibition. But while being in the same physical space implies remaining tied to its physicality, evolving within an additional virtual world goes beyond what is possible in the real world. Should AR designers perceive the same augmentations for the same physical objects? How could they create, organize and compare simultaneously different 3D virtual artifacts?

In this dissertation, I explore these questions with a conceptual framework that allows multiple versions of augmentations of the same physical object to coexist in parallel virtual spaces. According to this framework, AR designers can partially or totally desynchronize their virtual environment to generate their own content and then explore alternatives created by others. I illustrate it with a scenario of collaborative design in Augmented Reality, where two designers sketch in 3D around a physical

sewing mannequin to design a female dress.

On another hand, as we have seen during the COVID-19 pandemic, physical proximity can be challenged by an imposed distance. Still, designers need to collaborate and share digital and physical content. While video-mediated communication has been popular for remote collaboration via traditional video conferencing tools, this approach presents many limitations. First, it mainly relies on a single-camera viewpoint, which constrains the perception of the remote workspace. In particular, users do not have the freedom to inspect the remote workspace since cameras are commonly fixed in space or their positioning is out of their control. Also, pointing and annotating 3D content remains challenging from a 2D screen. Previous HCI research considered going beyond video-mediated tools with 3D workspace reconstructions. However, as these techniques require a high acquisition quality, I argue that they rely on complex systems and a high network bandwidth consumption.

Instead, I investigate solutions that rely on the use of complementary representations and views, which can be either fully virtual or augmented. I first consider the tradeoffs of three remote representations of an AR workspace : a fully virtual representation, a first-person view, and an external view. I then introduce a multi-view video-mediated communication system that combines these representations through interactive tools for navigation, previewing, pointing, and annotation. I report on a second user study that observed how 12 participants used ARgus to provide remote instructions for an AR furniture arrangement task.

REMERCIEMENTS

Cette thèse est le résultat d'un travail qui n'aurait pas pu être possible sans toutes les personnes qui m'ont aidé, accompagné et soutenu.

Tout d'abord, Theophanis Tsandilas et Cédric Fleury, sans qui rien de ces trois ans n'aurait été possibles. Vous m'avez montré comment se poser les bonnes questions, avec une disponibilité et une patience dont je vous serais éternellement reconnaissant.

Je remercie sincèrement les membres de mon jury de thèse, Anastasia Bezerianos, Ferran Argelaguet, Géry Casiez et Morten Fjeld pour votre temps, votre attention, vos questions et ces échanges particulièrement enrichissant que nous avons pu avoir pendant et après ma soutenance.

Malgré la pandémie de COVID-21, la vie de l'équipe Ex-Situ a su passer outre les confinements et s'avérer un soutien important. Premièrement, je souhaiterai remercier Wendy Mackay et Michel Beaudouin-Lafon, pour partager avec autant d'énergie et de bienveillance votre passion pour la recherche en Interaction Humain Machine. Merci également à Sarah Fdili Alaoui, Janin Koch et Nicolas Taffin pour faire vivre cette recherche avec conviction et incarner des modèles de carrière.

Mais ces années n'auraient pas été aussi vivantes sans la proximité de toutes celles et ceux avec qui elles ont été vécues. Merci Alexandre B., pour avoir partagé en plus d'un bureau toute l'intensité de la vie d'un thésard, des fous rires aux moments les plus difficiles. Merci Jean-Philippe et Antoine, pour votre soutien, qu'il soit au bout d'une corde ou au détour d'une pinte. Merci Junhang pour ces sessions sportives et culturelles, Martin pour l'intensité de nos échanges, Camille, Romane, Alexandre K. et Wissal pour toutes ces discussions enrichissantes et ces occasions où il a fallu représenter l'équipe en conférence, Manon, Anna, Capucine et Tove pour ces moments autour d'un tableau ou de gâteaux qui resteront dans ma mémoire. Merci à celles et ceux qui ont montré l'exemple, Miguel pour ton optimisme, Théo pour ton regard sur le monde, Liz pour tes attentions, Han pour ton enthousiasme et Viktor pour ta détermination. Enfin, je remercie celles et ceux dont nos routes se sont croisées durant ces trois ans, Léa, Baptiste, Eya, Bastien, Sally, Abby, Yi, Johnny, Benjamin et Julien.

Je souhaite également remercier Olivier Gladin pour ces discussions techniques et son aide dans le passage du théorique à la pratique, ainsi que Tifanie Bouchara, pour ces derniers conseils durant ma rédaction.

De même, j'adresse un remerciement tout particulier à mes professeurs de l'IUT de Rodez, pour avoir été des sources d'inspiration en plus de m'avoir ouvert les portes du monde de l'informatique.

Mais cette thèse s'est également construite en dehors du monde académique. Merci à Titouan, Ophélie, Malo, Marion, Camille, Karen, Julia, Philémon, Théo, Raphaël, Tom et Dan pour toutes vos attentions et ces moments hors du temps que nous avons pu partager. Vous m'avez donné autant que vous m'avez inspiré. Merci à Thibault, Morgane, Guillaume, José, Djibril et Hanna pour avoir été là du début à la fin et pour ce que nous avons pu vivre ensemble. Enfin, merci mille fois à Sitoz, Fonfon, LGF, JAMvep, Limi et Flo, pour votre amitié sans faille. Vous m'avez encore une fois prouvé que le temps et la distance ne sont que des détails. J'ai également une pensée spéciale pour la Bo213 et toute sa fraternité.

Je voudrai aussi remercier toute ma famille. Merci papa, maman, pour m'avoir encouragé avec autant d'amour. Merci Ava pour ton soutien et être toi, ce qui reste la plus belle leçon d'accomplissement.

Merci également à Michelle et Olivier Sablé, pour vos encouragements et pour avoir tenu à assister à ma soutenance

Et pour finir, merci infiniment à Chloé, pour ton amour, ta patience et ton soutien. Je n'aurai jamais assez de mot pour te montrer toute ma gratitude.

CONTENTS

1	Introduction	1
1.1	Terminology	2
1.2	Problem Formulation	2
1.3	Methodology and contributions	4
1.4	Thesis overview	5
2	Collaborative design in Mixed Reality	7
2.1	Computer mediated collaborative design	7
2.1.1	Design practices	8
2.1.2	Collaborative systems for design	12
2.1.3	Conclusion	17
2.2	Collaboration in Mixed Reality	18
2.2.1	The mechanics of collaboration	19
2.2.2	Co-located collaboration	22
2.2.3	Remote collaboration	28
2.2.4	Conclusion	33
2.3	Beyond physics	33
2.3.1	From physical to virtual spaces and objects	34
2.3.2	Reconstructing a remote MR workspace	39
2.3.3	View transition techniques	42
2.3.4	Conclusion	43
2.4	Positioning my work	44
3	Co-located collaboration with parallel versions	46
3.1	Scenario	47
3.1.1	Equipment	47
3.1.2	Context	47
3.1.3	Simultaneous co-creation	47
3.1.4	Divergence	50
3.1.5	Convergence	52
3.1.6	Summary of the scenario	53
3.2	Conceptual framework	53
3.2.1	Reifying versions	55
3.2.2	Browsing versions	56
3.2.3	Collaborating through Version Objects	59
3.3	Implementation	63
3.3.1	Interaction with the virtual content	63
3.3.2	Network communication between collaborators	64
3.4	Discussion	64
3.4.1	Scope of the framework	64
3.4.2	Understanding the framework	64
3.4.3	Implications for the design process	65
3.4.4	Extending the framework to Virtual Reality	65

3.4.5	Extending the versioning mechanisms to mixed-reality objects	66
4	Collaboration beyond physical proximity	67
4.1	Design problem	69
4.2	User study 1	70
4.2.1	Participants	71
4.2.2	Apparatus	73
4.2.3	Task	73
4.2.4	Design	75
4.2.5	Procedure	75
4.2.6	Data Collection and Measures	75
4.2.7	Results	76
4.2.8	Discussion	80
4.3	ARgus: A Multi-View Collaboration System	81
4.3.1	Combining Multiple Views	82
4.3.2	Supporting Navigation	82
4.3.3	Facilitating Communication	84
4.3.4	Architecture and Implementation	86
4.4	User Study 2	87
4.4.1	Participants	87
4.4.2	Apparatus and Conditions	87
4.4.3	Task	88
4.4.4	Design and Procedure	90
4.4.5	Data Collection and Measures	90
4.4.6	Questions and Hypotheses	91
4.4.7	Results	92
4.4.8	Discussion	96
5	Conclusion and Future Directions	99
5.1	Limitations	100
5.1.1	Beyond a unique space	100
5.1.2	Beyond a single viewpoint	102
5.2	Future work	102
	Bibliography	105

LIST OF FIGURES

Figure 2.1	Example of sketches in representative phases of a professional design process. Figure from Bousseau et al. [18].	9
Figure 2.2	3D sketch in AR made with a smartphone. Figure from Kwan and Fu [135].	10
Figure 2.3	Thinking processes of design teams. Process 1 (a) and process 2 (b). Adaptation from Stempfle and Badke-Schaub [221].	12
Figure 2.4	Traditional cognitive model of collaboration. Adaptation from Vera et al. [243].	13
Figure 2.5	Three designers using TEAM STORM to explore new form factors of a mobile device in public and private workspaces. Figure from Hailpern et al. [92].	15
Figure 2.6	A couple designing together a living room using AR. Figure from Shin et al. [208].	17
Figure 2.7	Virtual shadows of a collaborator’s arms are displayed on the surface of a table. Figure from Sakong and Nam [197].	20
Figure 2.8	Gaze and field-of-view representation of a remote collaborator. A virtual arrow indicates the position of the collaborator’s head in the workspace. Figure from Bai et al. [8].	21
Figure 2.9	Students interacting with virtual geometric shapes in AR. Figures from Kaufmann et al. [118].	23
Figure 2.10	Four users looking at a common 3D model placed on a rotating “cake platter”. Figure from Regenbrecht et al. [188].	27
Figure 2.11	Two AR collaborators working respectively from a) on-site and b) remote. Figure from Zhang et al. [262].	29
Figure 2.12	An agent (left) perceives virtual replicas of the arms of a remote expert (right) for a motherboard assembly task. Figure from Le Chenechal et al. [36].	32
Figure 2.13	A flexible and tangible Magic Lens used to investigate the internal components of an electronic appliance. Figure from Looser et al. [152].	36
Figure 2.14	Geometrical deformation of a virtual city model (a) to extend the visibility of occluded streets and buildings (b). Figure from Chen et al. [35].	37

- Figure 2.15 (a) An indoor user pointing at a location on a tabletop surface, which contains a virtual representation of the outdoor world. (b) Outdoor AR view showing the indoor user's hand appearing from the sky and pointing toward a location in the distance. Figure from Stafford et al. [220]. 39
- Figure 2.16 The Virtualized Reality studio: conceptual (a) and technical implementation (b). Figure from Kanade et al. [113]. 40
- Figure 2.17 The JackIn Space telepresence concept: (1) A remote user perceives A's first-person view (view 1). (2) The remote user's viewpoint smoothly moves out from the viewpoint of A to become a third-person perspective (view 2). (3) The user can freely change his/her view position, and alternatively, see from external cameras. (4) The user's viewpoint moves to user B's viewpoint (view 3). Figure from Komiyama et al. [129]. 43
- Figure 3.1 **Co-creation with Augmented Reality.** Gabrielle and Karl, two fashion designers, use AR headsets to design a virtual 3D model of a jacket around a physical sewing mannequin. 48
- Figure 3.2 **3D sketching tools** for the realization of a virtual 3D sketch on a physical sewing mannequin. (a) Tool for drawing virtual ribbons and (b) tool for erasing virtual ribbons. 49
- Figure 3.3 **Creation of a Version Object** to save a version of a 3D drawing. (a) Extraction of a Version Object by entering the 3D drawing, (b) moving the Virtual Object in the workspace and (c) storing the Virtual Object in the workspace. 50
- Figure 3.4 **Spherical 3D portal generated from a Version Object.** (a) 3D portal to the virtual space of the Version Object, (b) 3D sketch visible in the virtual space of the Version Object and (c) 3D sketch from the virtual space hidden by the 3D portal 54
- Figure 3.5 **Version Objects creation.** On the left, a version of a 3D sketch overlaying a physical sewing mannequin. In the middle, the corresponding Version Object. On the left, a collection of Version Objects. 55

- Figure 3.6 **Version Object visibility.** A shared Version Object is visible by all users. A non-shared Version Object is only visible by its creator. 56
- Figure 3.7 State diagram of virtual spaces. 56
- Figure 3.8 Preview (a) and 3D portal shaped as a rectangular parallelepiped (b) to simultaneously visualize two version of a 3D sketching on a physical sewing mannequin. In (a), the 3D sketching in 1 is a preview of the version contained in the Version Objects in 2 and 3. **Pink (A)** corresponds to the virtual elements shared by versions 2 and 3, **yellow (B)** to the virtual elements of version 2 not present in version 3 and **blue (C)** to the virtual elements of version 3 absent from version 2. In (b), the 3D sketching in 1, outside the portal, corresponds to the Version Object in 3. The 3D sketching in 2, inside the portal, corresponds to the Version Object in 4. 58
- Figure 3.9 State diagram of virtual spaces. 59
- Figure 3.10 **Desynchronized virtual spaces.** Collaborators' virtual spaces are in different states (a and d). They see the same physical mannequin but a different virtual 3D design (b and c). 60
- Figure 3.11 **Synchronized virtual spaces.** Collaborators' virtual spaces are in the same state (a and c). The collaborators have access to the same physical mannequin and the same 3D sketches (b). 61
- Figure 3.12 State diagram that describes the synchronization and desynchronization of collaborators' virtual space. 62
- Figure 3.13 System's architecture 63
- Figure 4.1 Remote-view configurations tested by our first study: HEADSET VIEW (left), EXTERNAL VIEW (middle) and VIRTUAL VIEW (right). A remote participant gives oral instructions to the AR user on how to position 3D shapes on a virtual support. 72

- Figure 4.2 (a) Remote participant interface used for our first study: tested view configuration on the left (VIRTUAL VIEW in this example) and website used to give instructions on the right. (b) Close-up of the website showing the target pattern: the UI widget on the right allows participants to rotate the pattern image. (c) Zoom-in on the AR user workspace showing the virtual board with the finalized task: colored axes help participants make the correspondence between the pattern on the image and the virtual board shown on the view. 74
- Figure 4.3 Comparing the perceived difficulty of different subtasks among configurations. For our analysis, we use Bayesian ordinal (cumulative probit) models [25], which map the original ordinal scale of Likert items to a *latent* continuous variable. The bars in the graph represent 95% *credible intervals* of mean differences over this continuous variable and can be treated as estimates of standardized effect sizes. Note that the unit of these differences is the standard deviation (SD) of the distribution of the latent variable. 77
- Figure 4.4 Distribution of participants' preferences: HEADSET VIEW vs. EXTERNAL VIEW (left) and HEADSET VIEW vs. VIRTUAL VIEW (right). 78
- Figure 4.5 Desktop interface of ARgus used by a remote collaborator for the redesign of a bicycle saddle. 82
- Figure 4.6 The remote user hovers the mouse over the headset of the 3D avatar (left) and Kinect 3D model (right) to display the preview of the HEADSET VIEW and the EXTERNAL VIEW respectively. 83
- Figure 4.7 Tools available in ARgus: SPHERICAL VIEW (a), VIRTUAL STICK (b) and annotations (c). 85
- Figure 4.8 System's architecture and implementation 86
- Figure 4.9 Experimenter wearing a Microsoft HoloLens 2 (a) collaborating with a participant using ARgus on a desktop computer (b). 88
- Figure 4.10 Symbols on the walls of the virtual house model to constrain the placements. 89
- Figure 4.11 Miniature furniture manipulated by the experimenter according to the participant's instructions. 90

- Figure 4.12 Use of the three view representations, the pointing stick, and the spherical view by the participants of the evaluation study for the main task under ARGUS. Circled participants were exposed to ARGUS first. 93
- Figure 4.13 Comparing the perceived efficiency of the two user interface configurations ($N = 12$). We use again Bayesian ordinal (cumulative probit) models [25]. The bars in the graph represent 95% credible intervals of mean differences over a latent continuous variable and can be treated as estimates of standardized effect sizes. 95
- Figure 4.14 Comparing the importance of verbal communication for the two user interface configurations ($N = 12$). We use again Bayesian ordinal (cumulative probit) models [25]. The bars in the graph represent 95% credible intervals of mean differences over a latent continuous variable and can be treated as estimates of standardized effect sizes. 95
- Figure 4.15 **Results of transcript analysis.** We compare the number of words pronounced by the 12 participants to provide instructions. The grey boxes at the left show the total number of words with non-instructions. The error bars at the right represent 95% confidence intervals derived from the t -distribution. 97

ACRONYMS

HCI	Human Computer Interaction
CSCW	Computer-Supported Cooperative Work
AR	Augmented Reality
MR	Mixed Reality
VR	Virtual Reality

INTRODUCTION

When people collaborate, they exchange, discuss, and confront different ideas. Novices call for the expertise of collaborators when they are stuck on a task. Similarly, experts help their less experienced partners with a subject they know. And since interactive technologies can help people work together, computer-mediated collaboration has been a major topic of interest for research in Human-Computer Interaction (HCI).

In this dissertation, I focus on collaborative tasks for 3D design, such as product and urban design. Sharing ideas is often necessary for solving design problems. Designers commonly work in groups. They gather in the same space to communicate with each other and mutualize their expertise. In architectural design, for example, structural engineers and architects meet together to find joint solutions and plan the creation of a building. In HCI practice, interaction designers conduct participatory design sessions with end users to brainstorm with them or receive their feedback about design alternatives and system prototypes. Designers also collaborate to create new designs by combining their individual designs [208]. Studies on design tasks [51] suggest that designers build better relationships within a group and produce better quality results when they share multiple prototypes with their peers.

A design process includes a series of stages, from coming up with original ideas to building the final artifact. The earlier stages require designers to think about what they want to create. Being more open-ended, the goal is to encourage designers to be creative and explore the space of possible design solutions. Sketching is a privileged practice during these early design stages and serves to extend the designers' line of thought [81]. The volume of sketches developed throughout the design process by design teams is known to correlate positively with the quality of the final design outcome [216]. However, sketching is not the only avenue for designers to represent their thinking [259]. In product design, individuals who do not draw can clarify things by realizing 3D prototypes. Therefore, design activities partly rely on sketching but also on creating 3D artifacts as prototypes.

The objects of interest of 3D design practice exist in space. To get meaning ("what can be done with it" vs. "how did it get there"), they need context, which their surroundings and other interacting objects can give. Thus, when designers design a new product, they consider the spatial context in which this product will be. As stated by Bryan Lawson [140], *"the trick of good design is to get an appropriate amount*

of order to meet the needs of the context or situation" [140]. For example, architects might consider the environment around the construction they are designing. Among others, they will pay attention to the lighting, the shape of the ground, and how this new construction fits in the landscape. Likewise, a car interior designer must be aware of the car itself to think about what will fit inside. Will the surface of this mirror cover the whole rear glass? Will the minimum inclination of the driver's seat be enough for people whose legs are shorter than average? To answer all these questions, the designer needs to view the artifact placed in context in 3D space.

Augmented Reality (AR) technologies help designers create and manipulate virtual artifacts in context with the physical world. They extend the visualization capabilities of traditional desktops by letting users move around a 3D model as they would do around a physical object. AR can also expose digital creations in interaction with physical content. In the past, designers had to spend significant time building physical prototypes to test their designs in physical space. Thus, AR design tools let them quickly iterate over their solutions and save effort and time.

My dissertation investigates the new opportunities of AR technologies and their implications for interaction design with a focus on collaborative design tasks.

1.1 TERMINOLOGY

In Milgram's *reality-virtuality continuum* [162], Mixed Reality (MR) bridges the gap between Reality and Virtual Reality (VR), and AR is a subset of it. However, the terms AR and MR are often used interchangeably because their boundaries are not clearly defined. In my dissertation, I sometimes refer to MR systems as many of their aspects are relevant to AR. Similarly, because reality and virtuality are connected within Milgram's continuum, I also discuss research literature on VR systems, which often provide insights about the design of AR systems.

1.2 PROBLEM FORMULATION

Fully virtual environments pose few constraints on how a workspace is shared among collaborators and provide unlimited opportunities for interaction design. As Sutherland [224] wrote in *The Ultimate Display*: "there is no reason why the objects displayed by a computer have to follow the ordinary rules of physical reality with which we are familiar". Metaphors facilitate interaction design. They rely on users' previous experience to ease the learning of new systems and can be inspired by situations from the physical world [6]. For example, as in a physical environment, a first-person view of a virtual space can follow the

user's head movement. Similarly, collaborators in a virtual environment can perceive virtual objects in the same position. Previous work in Virtual Reality has also explored interactions independent from physical metaphors. For example, users in the same physical space can be immersed in different virtual spaces [248]. Virtual objects can be duplicated when manipulation conflict arises, allowing parallel manipulations for design exploration [257]. Likewise, modifying the spatial geometry of virtual spaces helps visualize occluded 3D models [56].

In AR environments, in contrast, the virtual content needs to preserve its connection with the physical world. And unlike virtual spaces, physical objects only exist in one state at a time and are common for everyone. Thus, designing interactive systems for AR is especially challenging, as we need to consider the constraints of the physical world carefully, but also take advantage of traditional communication codes that naturally emerge during our physical interactions with others, such as gestures and gaze.

But AR technologies also allow for pushing the limits of interaction design far beyond what is possible in the physical world. For example, collaborators in an AR environment may be provided with personalized views of the virtual space. They can also get access to alternative or richer representations that are optimal for different aspects of their task. My goal is to understand such possibilities better and develop new interaction mechanisms for AR that help design teams collaborate more effectively.

I consider the physical distance between collaborators as a key dimension that determines my research directions and solutions. Specifically, I distinguish between *co-located* collaboration, when designers evolve in the same space and are physically close to each other, and *remote* collaboration, where designers work remotely and do not share the same physical space.

Co-located collaboration. In many physical-world situations, more than one designer needs to collaborate together around the same physical space. If this collaboration takes place in an AR environment, all collaborators will perceive the physical space and the physical objects in the exact same way. But what about their virtual spaces and their virtual design creations? Should they also be shared? I argue that this solution would be problematic, especially during the early design stages when designers want to quickly explore a large number of ideas before they discuss them with others. Also, as the number of designs increases, several questions emerge: Where should they be placed in space? Who is their owner? How can we help multiple designers iterate on these designs? What tools do they need to merge ideas together? The first problem that I study is framed around these questions.

Remote collaboration. In other situations, designers have to work together remotely. For this collaboration to be effective, the designers must be aware of the physical and virtual workspace of each other. They also need to keep track of the actions of their collaborators and be able to communicate with traditional interaction modalities, such as voice, gaze, and gesture. Video conferencing tools can assist such collaboration scenarios and became extremely popular during the COVID-19 pandemic [258]. Video-mediated communication is easy to set up and relies on devices (e.g., webcams or smartphones) already available to many users.

Video-mediated collaboration presents many limitations, though. First, it relies on a single-camera viewpoint, which constrains the perception of the remote workspace. In particular, users do not have the freedom to inspect the remote workspace since cameras are commonly fixed in space or their positioning is out of their control. The second problem I study in this dissertation is how to improve video-mediated communication, focusing again on 3D design tasks.

To address this problem, the past HCI literature has investigated a range of technologies that go beyond traditional video-mediated communication by allowing users to see and interact with a remote workspace through richer representations and augmented views [9]. Although recent systems have also tried to provide users with reconstructed virtual representations of a distant workspace (e.g., through point clouds [8], light fields [163] or mixed methods [238]), I argue that such technologies are still expensive in bandwidth and hardware resources, while the quality of reconstructed images is often low. Instead, I investigate solutions that rely on the use of complementary representations and views, which can be either fully virtual or augmented. My goal is to understand their trade-offs better and explore new interaction possibilities when they are combined together.

1.3 METHODOLOGY AND CONTRIBUTIONS

To study the first problem, I developed a *conceptual framework*, focusing on how to help designers collaborate in the same physical space through parallel versions of virtual design objects. According to Rogers [193], a framework is “a set of interrelated concepts and/or a set of specific questions intended to inform a particular domain area.” Frameworks are considered as theoretical contributions [255]. Girouard et al. [78] argue that they are “the foundation of strong research” and identify three framework types: abstracting, designing, and building. The framework that I present in this dissertation is partly designing and partly building: it conceptualizes interaction concepts while also aiming to help implement new systems. I illustrate the framework with a collaborative design scenario. I also present a system that implements its key functionality.

To investigate the second problem, I developed *ARgus*, a multi-view remote collaboration system. I also conducted two user studies. The goal of the first study was to inform the design of *ARgus*. The goal of the second study was to evaluate the system.

In summary, the main contributions of this dissertation are:

THEORETICAL. It introduces a conceptual framework for co-located collaboration. According to this framework, the collaborators can synchronize and desynchronize their virtual spaces through interactive containers that I call *Version Objects*. As their name implies, these objects enable users to create and interactively manage versions of 3D content in their workspace. The framework provides tools for navigation and comparison among versions.

TECHNICAL. It presents a system that applies the concepts of the above framework to collaborative 3D sketching. Furthermore, it introduces *ARgus*, a multi-view remote collaboration system. *ARgus* helps remote desktop users collaborate with a local designer who wears an AR headset and interacts in 3D space. The system provides tools for effectively switching among different representations of the remote workspace, virtually navigating in this space while pointing and annotating 3D content.

EMPIRICAL. It presents the results of two user studies. The first study explores the trade-offs of different workspace representations, viewpoints, and navigation controls for remote desktop users. Its results inform the design of *ARgus*. The second study observes the use of *ARgus* and compares it with a common configuration, where remote users have only access to a first-person view of the remote AR workspace. Reflecting on its results, I discuss the strengths of multi-view collaboration but also the costs associated with the additional complexity of interaction.

1.4 THESIS OVERVIEW

Chapter 2 presents previous work on collaborative design, focusing on the impact and possible roles of immersive technologies. I start by reviewing design workflows and methods, either from the point of view of individual designers or groups of designers who work together. I then look at how computer-mediated technologies support these design workflows and methods, with particular attention to the use of video and AR. I also review AR and VR systems and techniques that support collaboration between co-located or remote users and identify collaboration mechanisms involved in this context. Finally, I examine alternative virtual representations of physical content (i.e., physical design artifacts, the physical surroundings, and the actual people who engage in collaboration), focusing on virtual navigation

and transformation techniques that have no equivalent in the real world.

Chapter 3 examines scenarios of co-located collaboration during the early phases of design when sketching is commonly used. I present AR techniques for generating and comparing alternative designs, which I frame within a conceptual framework. The framework supports interaction with parallel design versions, which can be shared or remain private. In my design, I take to leverage the benefits of physical interaction between collaborators, but I also try to go beyond its traditional norms or the constraints of real-world physics. I demonstrate my approach with a scenario in which two fashion designers sketch ideas of a virtual cloth around a physical sewing mannequin. I describe a system that implements this scenario. This work will be presented as long paper at the 34th Francophone Conference on Human-Computer Interaction (IHM'23) in April 2023 [60].

Chapter 4 focuses instead on remote collaboration. I study the situation in which a local AR worker interacts with a 3D scene that may combine both virtual and physical components. Yet, their remote collaborators do not have access to specialized AR equipment, so they interact instead through their laptops or tablets. As a first step, I identify three alternative workspace representations that visualize the local workspace through different viewpoints and provide different navigation controls: an augmented first-person view, an augmented external view, and a fully virtual view with independent navigation controls. I present the results of a user study (24 participants) that assesses the trade-offs of these workspace representations. I then introduce ARgus, a system that provides fluid transitions among these three views and enhances them with direct pointing and annotating tools. Finally, I present the results of a user study that observes how 12 participants use the system to provide remote instructions for a furniture arrangement task. The study further examines how ARgus helps remote users communicate instructions compared to a system that only offers an augmented first-person view. This work was presented at the ACM Conference on Computer-Supported Cooperative Work And Social Computing (CSCW) in November 2022 [59] and demonstrated at the 33rd Francophone Conference on Human-Computer Interaction (IHM'22) in 2022 [58].

Chapter 5 concludes this dissertation. I summarize the contributions of my work in the light of related work in the HCI literature. Finally, I discuss the limitations of this work and suggest opportunities for further investigations.

This chapter explores collaboration in mixed reality (MR) systems. Its focus is on activities and methods used in collaborative design and how computer-mediated systems aid these design practices. I differentiate between co-located and remote collaboration settings. For co-located settings, I review alternative material setups and discuss their impact on collaboration. I also examine issues linked to how collaborators organize their workspace, position themselves, and communicate with each other through gestures. For remote settings, I identify techniques that enable distant collaborators to share their workspace and establish spatial references for communication. I give particular attention to the role of video as the primary medium of remote collaboration and discuss its use in the context of MR systems. Furthermore, I discuss techniques that aim to support users' awareness of their collaborators' actions and their workspace and mechanisms that help them divide and manage their private and shared space.

My analysis in this chapter emphasizes aspects of a virtual world that do not have direct analogies with the physical world. In particular, I review transformations and geometry deformation techniques, approaches that play with the scale of collaborators and their environments, mixed representations, and view transition techniques.

I end this chapter with a discussion of the limitations of existing MR collaboration systems. I identify crucial research gaps, which I then use to motivate the work presented in this dissertation.

2.1 COMPUTER MEDIATED COLLABORATIVE DESIGN

Design is a matter of finding and solving problems [140]. It is an activity that involves the selection, combination, and creation of items to produce ideas or physical objects. According to Sydney Gregory [84], *"the process of design [is] a process [whose] the pattern of which is the same whether it deals with the design of a new oil refinery, the construction of a cathedral, or the writing of Dante's Divine Comedy"*. Design is a multidisciplinary activity and often involves people with diverse skills, such as engineers, artists, architects, and fashion designers.

Collaboration is an important aspect of design. Because designers do not share the same knowledge or perspective, they need to mutualize their expertise through cooperation, coordination, and collaboration [158]. By doing so, they address cooperative processes, such as the establishment of a common ground, perspective clarification, and convergence mechanisms [46].

Computers have the potential to support some of our cognitive and creative processes which are fundamental aspects of design [141]. In this section, I further explore how computer-mediated technology supports collaborative design. I first present the theoretical aspects of the design process, from an individual to a collaborative perspective. I then introduce systems that support co-design, taking a closer look at those relying on video-mediated communication and MR.

2.1.1 Design practices

Design is a key process in many fields, particularly in architecture, engineering, or construction. It involves creating, manipulating and managing information [141] to generate artifacts. But, even though design activities mainly rely on “transforming, combining and adapting elements of previous designs, as well as elements and aspects of other objects, images and phenomena” [53], it also depends on the creation of new ideas. This is consistent with the work of De Bono [44], who introduces the concept of *lateral thinking*, in opposition to the *vertical thinking*. Lateral thinking aims to explore different ideas in parallel, while vertical thinking concentrates on exploring one idea in depth. An idea can be modified in many ways during the design phase. Lawson [140] makes a distinction between the *evolutionary modifications*, which consist in adding gradual modifications of an idea, and the *revolutionary modifications*, aiming to start from a new idea.

A generic model of design activity has been proposed by Stempfle et al. [221]. While this model is valid for any type of thinking and problem-solving, they propose four basic cognitive operations to address design thinking: generation, exploration, comparison, and selection.

The first two serve to extend the problem space, while the last two restrict it. This echoes the work of Rittel et al. [190], which proposes to see the creative process as a combination of divergent and convergent processes.

2.1.1.1 Sketching in design

According to Donald Schön [205], sketches are part of the mental process of designing. This makes sense with Lawson’s work [140], who observed that sketching is at the center of most of the designing processes. Sketches are important for designers as they allow them to *ideate* and solve problems. It also provides a way to communicate with other designers [30].

As the designing process can be seen as an individual discussion for the designer, Schön suggests that a designer “has a conversation with a drawing.” The act of sketching becomes thus the continuity of a line of thought. Goel [81] presents sketches like a symbol system that is imprecise, ambiguous, fluid, amorphous, and indeterminate. This

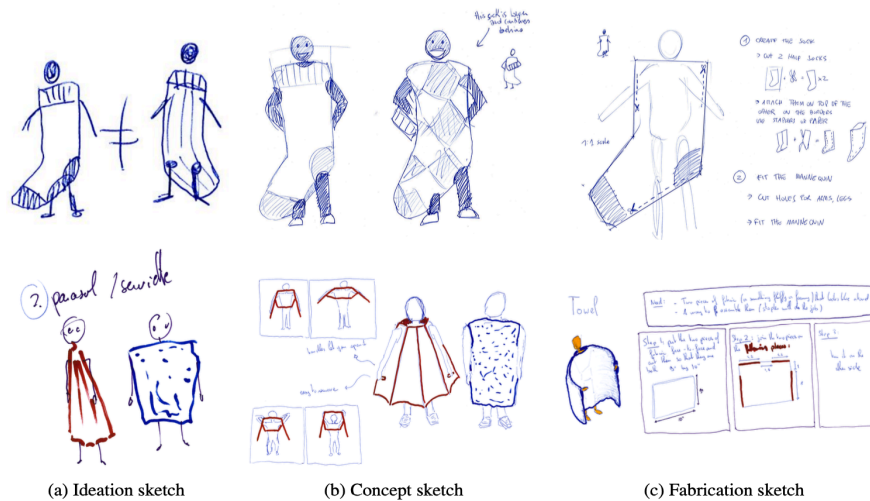


Figure 2.1: Example of sketches in representative phases of a professional design process. Figure from Bousseau et al. [18].

kind of system allows association mechanisms that ease divergence and widen the problem space.

Sketches are used in different stages of the design process [55] [93]. There is a distinction among ideation sketches, conceptual sketches, sketches used for presentation purposes, and sketches for fabrication. Examples of such sketches are presented in Figure 2.1.

Sketching with physical materials has limitations. For example, Damm et al. [40] show that modifying a drawing on a whiteboard can be laborious. The fixed space of the board is limited, and it is impossible to save or reload a drawing. Computer tools for sketching can overcome such limitations and extend our cognitive and creative processes [141]. Digital sketches are "easy to edit, store, duplicate, modify, and search" [136]. Also, computer drawing systems can provide more advanced assistance for 3D modeling. For example, *DreamSketch* can transform free-form 2D sketches into 3D models and ease the exploration of design ideas with generative algorithms [120].

2.1.1.2 Sketching with Mixed Reality

With *Spacedesign* [63], designers realize free-form curves and surfaces using MR. They can perform freehand sketching, surfacing, and engineering visualization using a configurable semi-immersive workbench-like workspace. This system materializes ideas as 3D content and realizes pre-production physical prototypes using additional shapes, textures, and annotations.

Different techniques exist to realize three-dimensional sketching in AR. Dudley et al. [52] studied three of them: freehand, tapline, and gogo-tapline. *Freehand* consists in continuously drawing using hand movements. This technique is fast, although not necessarily accurate.

Tapline draws lines by specifying control points and tries to balance between speed and accuracy. Finally, the GoGo-Tapline technique relies on Tapline, but instead of using the position of the index finger to specify control points, it uses a virtual cursor that moves relative to the user's arm extension. The authors showed that each technique has its own trade-off and that switching among them is the best way to achieve fluid sketching in AR.

Sketching can also be performed within a fully virtual environment. *SketchTab3D* [16] is an example of 3D sketch generation using a tablet in an immersive virtual environment. It allows users to move, duplicate, and scale sketches. Likewise, *Hyve-3D* [49] relies on an immersive system that enables users to sketch virtual content in a virtual 3D space. Its users can view and manipulate 3D content in immersive environments but also on conventional 2D displays and projections. With *DesignAR* [189], users can create and manipulate 3D objects from sketch input on an interactive 2D surface using an AR headset. A key limitation of these systems is that they do not accommodate physical-world objects.

By integrating a virtual space into the physical world, AR can help designers to take advantage of their physical environment as they explore ideas. For example, using spatial augmented reality with a projector and a depth camera, designers can draw virtual sketches on top of physical content [139]. They can also use a tablet [115] or a smartphone [135] (see Figure 2.2) to draw and place 3D sketches in their surroundings. Similarly, designers can use a tablet to scan a place and sketch in 3D using a projection from the screen [99]. These sketches can then be used to generate 3D volumes.



Figure 2.2: 3D sketch in AR made with a smartphone. Figure from Kwan and Fu [135].

2.1.1.3 Collaborative design

In collaborative design, several design partners are working on a solution together [42]. They share the same goal and contribute to its achievement with their specific competencies, and they create new designs by combining their individual designs [208]. Dow et al. [51] found that designers create higher-quality work when they produce and share multiple alternatives of their work with their collaborators, as they explore more different ideas, integrate more of their partners' features, and engage in more productive design conversations.

However, collaborative design can be hindered by the designers' inability to understand and appreciate the views of their collaborators, especially those from different disciplines [17]. Stempfle and Badke-Schaub [221] highlight the fact that to solve a design problem, design teams need to structure and organize the group process in addition to dealing with the design task. They raise two focuses of actions: the content and the process. The *content* is defined by the following steps:

1. **Goal clarification:** communicative acts dealing with the goal space.
2. **Solution generation:** proposals and solution ideas concerning the design task.
3. **Analysis:** questions and answers concerning the solution space.
4. **Evaluation:** positive and negative evaluations concerning the solution space.
5. **Decision:** decisions for or against a solution idea.
6. **Control:** control of the implementation of a solution idea.

The *process* is defined by these five steps:

1. **Planning:** proposals concerning the group process (how to proceed, how to distribute tasks, etc.).
2. **Analysis:** questions and answers concerning the group process.
3. **Evaluation:** positive and negative evaluations of the group process.
4. **Decision:** decisions concerning the group process.
5. **Control:** summary or control of group members' work.

Stempfle and Badke-Schaub also provide a two-process-thinking theory in design teams [221]. The first process, called *process 1*, is described in Figure 2.3a.

This process closely matches the "natural" thinking process of design teams. It results in considerable savings in time and cognitive effort spent on a problem. However, it tends to produce errors when

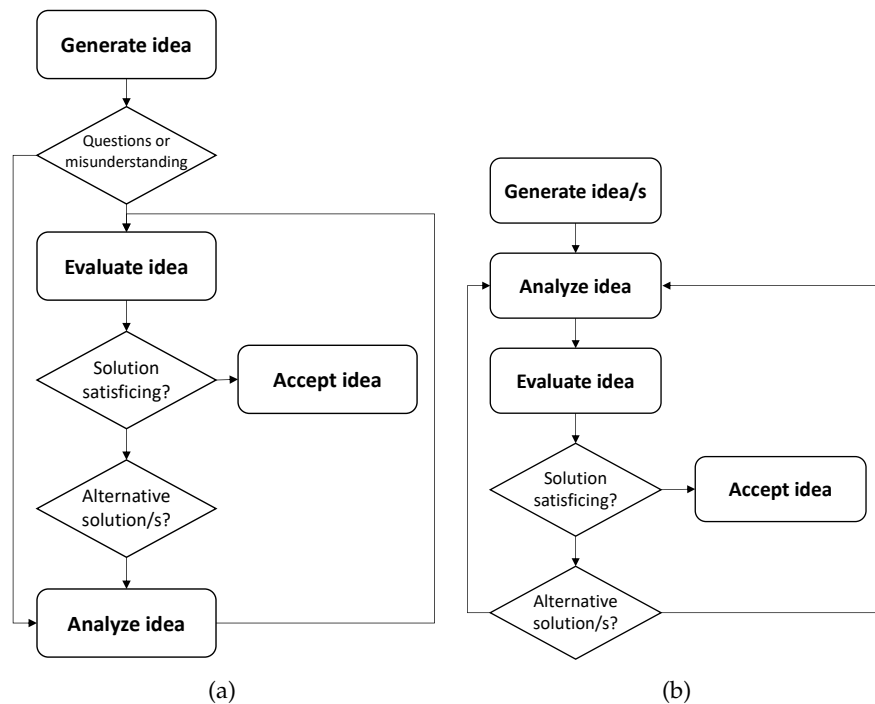


Figure 2.3: **Thinking processes of design teams.** Process 1 (a) and process 2 (b). Adaptation from Stempfle and Badke-Schaub [221].

the complexity of the design problem increase. Design teams then switch to another process, called *process 2* (see Figure 2.3b).

This second process is characterized by solution ideas being followed by analysis. Whereas it yields better qualitative results for complex problems, it does take more time and greater effort than the first process. Design teams will naturally tend to employ the latter.

With the hypothesis that expert collaborative problem-solving is structured very much like individual expert problem-solving, Vera et al. [243] propose a traditional cognitive model of collaboration (see Figure 2.4). First, a *meta-planning* process breaks down the problem into individually manageable units as well as about how and when the collaborators should come together to integrate their individual efforts. Then, a *negotiation* process occurs, which involves specific aspects of the design problem. Following this step, each collaborator separately engages in well-learned routine expert problem-solving guided by the meta-plan previously agreed. Finally, collaborators evaluate the outcome and finish or iterate these steps.

2.1.2 Collaborative systems for design

Groupware supports collaborative tasks, helping designers reach their collective design goals. This term covers both the intentional group processes and the software that supports them [108]. Groupware supports co-design when designers cannot meet in the same place

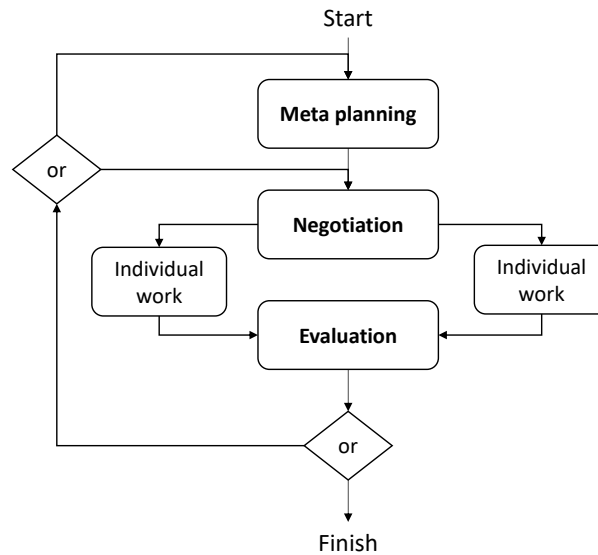


Figure 2.4: **Traditional cognitive model of collaboration.** Adaptation from Vera et al. [243].

at the same time. Collaborative drawing is, for example, supported by systems, such as Gambit [199], Calico [149], or Sketchboard [270]. Other systems like *TelePICTIVE* [211] support the design of Graphical User Interfaces by relying on metaphors of common office objects, such as Post-it notes. Such systems have been studied for various application domains, e.g., for design evaluation in architecture [94]. But Kirk et al. [125] also observed that showing the collaborators' hands as video in addition to the digital sketch can improve collaboration performance.

Collaborative drawing has been made possible using video-based communication with systems, such as *VideoDraw* [231]. *Face2Face* [54] also uses video to support remote collaboration by providing the illusion of a vertical glass between collaborators. Presence, eye contact, gaze awareness, and gestures are supported and provide interaction with virtual 2D content to support the collaborative design process.

Nishino et al. [167] propose instead a Distributed Virtual Environment (DVE) to share a 3D model among designers. Likewise, *NetSketch* [138] provides a shared virtual space for distributed designers to create, manipulate and view 3D models. The system enables the seamless creation of local copies of objects for each collaborator and manages manipulation conflicts due to the absence of locking mechanisms. This strategy was later used in the context of MR collaboration, e.g., in *Spacetime* [257]. *NetSketch* also shares viewpoints between participants using a "picture-in-picture" on collaborators' screens. In other work [157], real-time updates of a shared 3D scene can transform an iterative design process into a real-time design process. With

Multiuser Groups for Conceptual Understanding and Prototyping, Cera et al. [28] provide a shared multimodal 3D workspace where remote designers can collaboratively sketch conceptual designs and author design semantics.

Finally, Lee et al. [144] observe that collaborators spend more time on systems that support sketching or viewing of 3D objects when they are distant than when they are face-to-face.

2.1.2.1 *Design alternatives*

The most intense collaboration takes place in the early phases of the development activity, such as concept generation and creation of design alternatives [4]. The ability for designers to observe the work of their collaborators during a group design session is key to increasing user's overall creativity and satisfaction with their final sketch [247]. *FlexiSketchTeam* [256] allows multiple collaborators to simultaneously draw on the same document with tablets or electronic whiteboards. Conflicts are managed via a locking system. Sketched objects are reusable and accessible in the form of categories to all collaborators. Piya et al. [182] go even further with *Co-3deator*, a collaborative system for creating 3D models from tablets. These models are shared to be reused or modified by collaborators. Likewise, *Skwiki* [263] proposes a content creation framework focusing on the creation and modification of alternatives. Generation and exploration of design alternatives are also possible with wall-sized displays [169]. Such systems aim to foster the creation and reduce iterations during industrial design reviews.

However, *TEAM STORM* [92] is among the first collaborative systems supporting parallel sketching on various devices. It proposes public and private workspaces to share or not some sketching among collaborators (see Figure 2.5). Geyer et al. [77] further explore this approach with *IdeaVis*. This system supports ideation through collaborative parallel sketching but using traditional supports, such as paper sheets or pencils. *DigiMetaplan* [142] extends these mechanisms by providing versioning features and supporting both remote and co-located collaboration.

2.1.2.2 *Version management*

Keeping track of design artifacts through time helps designers retrieve and organize their creations. Version management systems have first been explored in design with CAD models. Katz et al. [117] suggest a data model for organizing a design database across time. They define design objects which aggregate design information and connect them using version, configuration, and equivalence relationships. This model is implemented using a version server with workspaces, check-in and check-out actions, currency within version histories, and dynamic configurations through layers and contexts. Wagner et

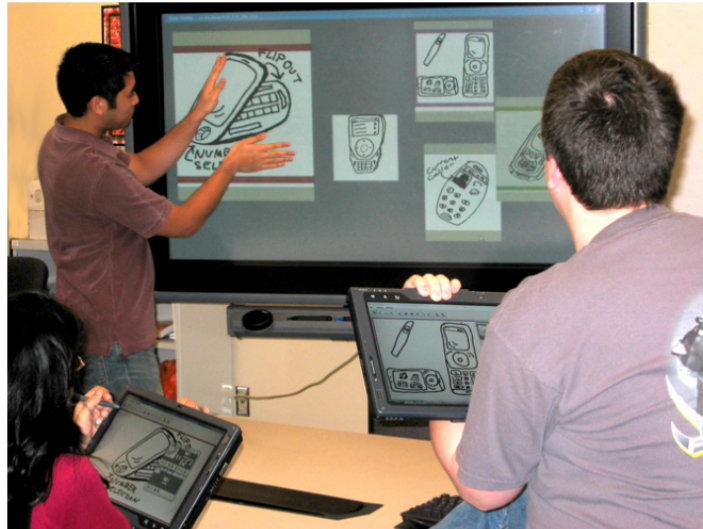


Figure 2.5: Three designers using TEAM STORM to explore new form factors of a mobile device in public and private workspaces. Figure from Hailpern et al. [92].

al. [246] consider design objects versioning through the GARDEN framework. They provide a data model supporting representations of complex design objects and design version management along three dimensions following the design evolution: alternatives, views, and revisions. Inspired by textual versioning in software engineering practices, Cristie and Joyce [39] explored parametric design exploration with *GHShot*, a versioning tool for parametric modeling. They show that versioning can aid design exploration, helps the documentation process, and enhances design sharing.

2.1.2.3 Video-mediated collaboration

When people cannot work in the same place, video remains one of the most common communication mediums. The goal of a video-mediated collaboration is to bring a common ground of understanding (or *conversational grounding* [67]) and support *workspace awareness* [89], or a “*shared person space*” that includes “*facial expressions, voice, gaze and body language*” [24]. Visualizing the activity is also useful. As Tang et al. [230] report, the process of creating and using drawings conveys information that is not contained in the resulting drawings [230]. Gergle et al. [76] also observe that collaborators adapt their communication to the presence or absence of a shared visual space. They found that action replaces explicit verbal instruction when a shared view of the workspace is available.

Video helps to show understanding, forecast responses, communicate non-verbal information, enhance verbal descriptions, manage pauses, and express attitudes compared to audio only [102]. Video-mediated communication can then help to handle conflicts and inter-

action-intense activities. Jones et al. [109] also observed that camera manipulation replaces some verbal interaction during a collaboration. Though, video interaction induces difficulties in managing turn-taking, noticing small movements through peripheral vision, having side conversations, and pointing and manipulating objects in the collaborator's workspace [102].

Video-mediated communication does not produce the same conversation style as face-to-face interaction [95]. A reason is that video cannot adequately convey all the non-verbal cues available in face-to-face communication, introducing a functional seam between participants [103]. Nguyen et al. [166] also shows that video conferencing can significantly hinder the trust formation process and decrease investment in comparison with face-to-face meetings. In an audio-visual video conferencing with a shared whiteboard, the amount of time spent discussing design ideas seems to decrease compared to a face-to-face situation. This is due to an increased need for communication control, communication about technology, and social communication which is not the case without video [69].

Vera et al. [243] observe that changing the communication channel between chat line and video-conferencing in an architecture design problem has an effect on the profile of communication content. This one can change from a high-level design (including broad decisions which will affect significant aspects of later decisions) to a low-level design ratio (placing individual elements, drafting, resolving issues of drafting, etc.). However, the change in communication channels does not influence the design outcome.

2.1.2.4 Collaborative design in MR

Augmented Reality supports various collaborative design processes. Shin et al. [208] observed several collaboration behaviors in the design of living rooms by couples using AR tablets (see Figure 2.6). Couples manipulate objects by taking turns when they explore different ideas. This is not the case when they share a common vision of their design. Couples occupy the space differently, from moving around to staying in the same spot. They also apply different strategies, from implementing a final design based on an individual design to designing together all design steps. *ARCritique* [145] relies on a mobile AR application for collaboration around a physical object. It supports pointing and drawing and shows that spatial referencing interactions can make communication more efficient and fluid than videoconferencing tools during critique sessions.

Burkhardt et al. [23] use AR for architectural design with *EsQUIsE* and a "Virtual Desktop" [196], a sketching device relying on projected content on a desk. They find that the design process occurs whether collaborators are co-located with a Virtual Desktop or distant, using *EsQUIsE* on tablets with video-conferencing software. However, the



Figure 2.6: A couple designing together a living room using AR. Figure from Shin et al. [208].

lack of awareness in the remote situation leads to difficulties in sharing context. Using projected AR, *C-Space* [215] supports spatial design exploration by providing an interactive and collaborative platform. This system assists designers in the development of design ideas in the early design phase. It supports design history, design reference exploration with AR, and prototyping using tangible interfaces.

Spacetime [257] introduces a new set of interaction concepts for collaborative editing: Containers, Parallel Objects, and Avatar Objects. *Containers* objectify space and time and situate interaction in context. They store a selection of objects, serve as proxies, viewports, or portals, and provide structural as well as temporal information. *Parallel Objects* allow for design comparisons, while *Avatar Object* objectify the actual users, such that their collaborators can directly manipulate them to share their views. Despite their sophistication, the above mechanisms were designed for VR. Unfortunately, it is unclear whether or how the above concepts could apply to collaborative AR environments.

2.1.3 Conclusion

I examined how MR extends the creation space of designers beyond the computer screen and helps contextualize design artifacts. Furthermore, MR environments can respond to the limitations of traditional 2D visualization methods and enable users to interact with virtual content in three dimensions. However, many features of traditional collaborative design systems cannot be easily transferred to an MR context. In particular, we need to rethink how to split the physical and virtual workspace to allow for both shared and private spaces of design exploration and creation. We also need to develop techniques that deal with occlusions, avoid or handle conflicts among collaborators, and help them discuss alternatives to eventually diverge their solutions.

I also discussed how video-mediated communication supports collaboration when designers do not share the same physical space. Video is a cheap yet powerful medium that communicates “facial expressions, voice, gaze, and body language” [24]. Those communication methods are crucial for settling “conversational grounding” [67]. However, videos alone cannot convey all communication cues available in face-to-face collaboration. In particular, it does not preserve spatial references in the 3D workspace, and its content is not interactive. Yet, spatial references and interaction are both significant for design tasks. How to preserve as many of the strengths of video communication as possible while leveraging the capabilities of MR technologies is a promising research direction.

2.2 COLLABORATION IN MIXED REALITY

While the first devices to access and share digital information were restrained to the use of 2D interfaces, the emergence of 3D input and output devices has more recently shifted the HCI design perspectives. *Collaborative Virtual Environments*, as well as *Virtual Reality* and *Augmented Reality* technologies, allow their users to interact with digital content in three dimensions.

Because people do not all have the same skills, they have to collaborate to share their competencies. Collaboration is a key aspect in many fields, such as architectural design, construction, engineering, and health care. Information technology has been increasingly used to support collaboration in such domains. Within this context, the CSCW community has studied a variety of *groupware*, where groupware can be defined as “intentional group processes plus software to support them” [108]. According to Lynda M. Applegate [7], groupware can be categorized through time and space. Collaborators can interact at the same time or through a delayed interaction from the same physical place or different locations.

When multiple users work together in MR environment at the same time, we say that their collaboration is *synchronous*. But collaboration can also be *asynchronous*, i.e., collaborators can interact with the artifacts they create at different moments. MR devices can support asynchronous collaboration by keeping track of collaborators’ edits through their memory storage or decentralized storage via the Internet.

Collaboration can naturally happen when people gather in the same place. In this case, they share the same physical environment while their position and activity can be integrated into a virtual space. We refer to this type of collaboration as *co-located collaboration*. However, in many situations, collaborators cannot gather around the same place, e.g., due to geographic distances or unusual circumstances, such as a lockdown. In this case, we refer to a *remote collaboration*. In the

following subsections, I first describe the concept of *awareness*, before discussing co-located and remote collaboration. I make a separation between these two situations because although the same MR devices can be used, each one presents different constraints on how to support awareness. I also examine additional dimensions of a collaboration environment, such as the mobility of its users, its virtual content, the roles of its users, and the visualization hardware. Several of these dimensions are discussed by Brockmann et al. [19].

2.2.1 *The mechanics of collaboration*

Pinelle et al. [176] identify the *mechanics of collaboration* as actions and interactions that group members need to perform to complete a collaborative task. They classify these actions and interactions into two large categories: *communication* and *coordination*.

Communication regroups two main activities: explicit communication and information gathering. *Explicit communication* is driven by the mechanics of spoken, written, or gestural messages, as well as deictic references or manifesting actions. For example, actions of pointing, indicating, drawing, and conversing refer to this category. *Information gathering*, instead, focuses on the perception of a workspace: what happens to the environment and the people inside this environment.

Coordination refers to shared access to tools, objects, space, and time, as well as the transfer of objects. The former relies on mechanics, such as obtaining a resource, reserving a resource, or protecting a work. The latter focuses on giving or taking objects, verbally offering or accepting objects, as well as placing objects and notifications.

An important requirement for information-gathering activities is *workspace awareness*, while coordination depends on *territoriality* or on how collaborators occupy and manage their space. Below, I dive into these two concepts. I discuss their role in Mixed Reality applications through several examples.

2.2.1.1 *Workspace awareness*

In the CSCW field, awareness refers to "a person's being or becoming aware of something" [203]. To perform an effective collaboration, collaborators need a sense of awareness of the other participants, in addition to a shared view of a collaborative workspace [230]. Collaborators also need a sense of who is there, what is there, where, when, and how. To conceptualize these information requirements, Gutwin and Greenberg [89] define *workspace awareness* as "the up-to-the-moment understanding of another person's interaction with the shared workspace".

Supporting workspace awareness is a requirement for effective remote collaboration. Workspace awareness can help "recognize opportunities for closer coupling, reducing the effort needed for verbal



Figure 2.7: Virtual shadows of a collaborator's arms are displayed on the surface of a table. Figure from Sakong and Nam [197].

communication, simplifying coordination, allowing people to act in anticipation of others, and providing context for appropriate help and assistance" [89]. Similarly to a face-to-face situation, workspace awareness is maintained by a *perception-action cycle*. People perceive information from their environment, add it to previously gathered information, and then use it to obtain additional information about the workspace. For example, when users try to help their collaborators, they might seek visual evidence about their collaborators' understanding, especially when their common vocabulary or their past interactions do not provide enough information about how people understand and make use of this help [172].

The HCI literature has introduced a range of techniques to support awareness. Sakong et al. [197] used physical turntables whose rotation was synchronized to help users better perceive their collaborators' interaction. In addition, virtual shadows of the users' hands were projected on their collaborator's table to give them more detailed feedback about their manipulation actions (see Figure 2.7).

Other work has explored techniques for communicating the users' field of view and their gaze. Piumsomboon et al. [181] represent the gaze and field of view of collaborators in 3D as a pyramid. Bai et al. [8] also use a ray-cast line to represent gaze (see Figure 2.8). They observe that communicating gaze in combination with gestures can improve task completion time and reinforce the sense of co-presence among remote collaborators [8]. Lee et al. [143] highlight the collaborator's view with a colored rectangle and provide a virtual arrow that indicates the collaborator's view when it exits the field of view of the user. Workspace awareness can also be extended by increasing the field of view with distorted peripheral lenses or by exaggerating the representation of actions so that others can see them [96]. However, some additional cues might not be beneficial to users. For example, showing the collaborator's head gaze looming has no clear advantages on the user's perceptual load [29].

In addition to information about the direction and field of view of users' gaze, recent work [106, 107] has also looked at how to pro-

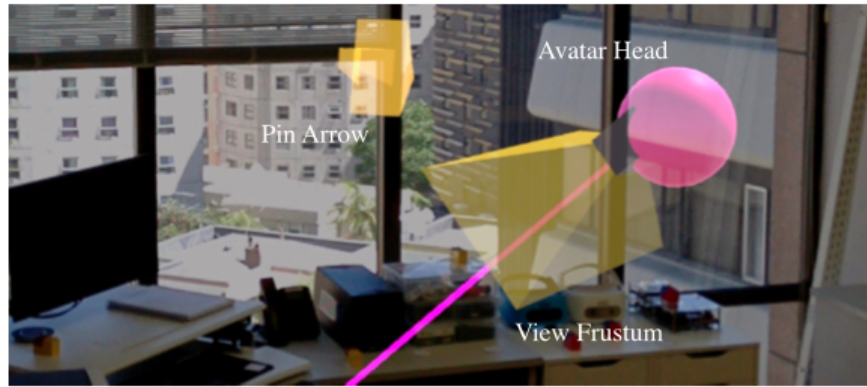


Figure 2.8: Gaze and field-of-view representation of a remote collaborator. A virtual arrow indicates the position of the collaborator’s head in the workspace. Figure from Bai et al. [8].

vide *behavioral cues*, such as information about whether users fixate their eyes on an object or whether their eyes blink or shift, information about their synchronized eye fixations, etc. Jing et al. [106] observe that the visualization of bi-directional gaze with behavioral cues makes collaborative tasks less physically demanding. It also makes collaborators more confident that their own gaze information is accurately delivered. It can also facilitate the identification and alignment of gaze states and behaviors, as opposed to commonly used unidirectional gaze representations with no behavioral cues [107].

Workspace awareness also depends on the modality of communication, and some previous results show that multi-modal solutions are especially beneficial. Wang et al. [249] observed that supporting both gestures and speech during a collaborative task led to faster performance than speech support only. Similarly, Kontogiorgos et al. [130] studied a collaborative furniture assembling task and found that verbal cues only are not enough for reliably identifying objects. In particular, they observed that gaze and head direction are useful for estimating the general placement of objects but not precise enough to help users decide on an exact target. They also noticed that combining information about gaze or head direction with speech increased classification accuracy, where information about the head direction seemed slightly more effective than eye gaze.

The role of gestures has been more extensively studied by Kirk et al. [124]. The authors identified gestures that promoted awareness in an MR ecology and categorized them through a corpus of “gestural phases.” They observe that hands can be used to establish agreement on the sharing of a common frame of reference (Flashing Hand), to coordinate the selection of correct objects (Wavering Hands), to indicate how objects should be assembled (Mimicking Hands), to show precise movement to align and make fit objects (Inhabited Hand), to

clarify instructions and repairs mistakes (Negating Hand), or to order turn-taking (Parked Hands).

2.2.1.2 Territoriality

Space is where people can gather together and collaborate. Space allows people to act [185]. Projecting energy and information of people into a space defines *territories* [214]. These territories have different sizes and can “range in size from chairs, seats, or sides of a table, to street blocks” [233]. *Territoriality* can be seen as “the activity that humans carry out in the space that is given or provided to them in common, within the limits of the conception that they have of it” [185]. Taylor [233] defines human territoriality as follows:

“an interlocking system of attitudes, sentiments, and behaviors that are specific to a particular, usually delimited, site or location, which, in the context of individuals, or a small group as a whole, reflect and reinforce, for those individuals or group some degree of excludability of use, responsibility for, and control over activities in these specific sites” [233].

Territoriality is exclusive to the one who occupies or defines territories [214]. Because delimited territories present different characteristics, territoriality can also be seen as a behavior related to the spatial organization of such territories. It mediates social interactions, considering who occupies which territory [5], to make collaboration possible.

Studying collaboration around tabletops, Scott et al. [206] distinguished among *personal*, *group*, and *storage* territories. *Personal territories* serve to reserve an area or task resources. These territories can provide a “safe” place where users can explore their own ideas before exposing them to collaborators [230]. *Group territories* are spaces allocated to the main task activities. Transfer of task resources among collaborators happens in this space. Finally, *storage territories* are used to store task resources and non-task items.

Men and Bryan-Kinns [161] studied the role of personal and group territories for social creative tasks in VR. They found that personal working spaces enable people to explore ideas without disturbing others and be more productive. However, reserving space for personal use can also lead to a smaller group territory and a greater distance among collaborators. While these results offer initial insights into personal and public spaces within virtual worlds, the topic of territoriality in collaborative Mixed Reality deserves further exploration.

2.2.2 Co-located collaboration

Augmented Reality has been used for a range of applications of co-located collaboration: learning [118] (see Figure 2.9), visualization of



Figure 2.9: Students interacting with virtual geometric shapes in AR. Figures from Kaufmann et al. [118].

scientific data [225], architectural design [20], sales [148], etc. Back in 1998, Schmalstieg et al. [225] implemented *Studierstube*, one of the first collaborative AR systems for the visualization of scientific data. Equipped with an AR headset, the collaborators could perceive 3D data representations that augmented physical objects. The authors defined five key properties of collaborative systems that use Augmented Reality:

- **Virtuality:** Users can access objects that are unavailable or non-existent in the physical world.
- **Augmentation:** Objects from the physical world can be augmented with virtual content that lies in 3D space.
- **Multi-user support:** Users can discuss with each other, design in parallel, and collaborate.
- **Independence:** Each user can move freely, independently of his or her collaborators.
- **Sharing vs. Individuality:** Different users may simultaneously see the same virtual content, or alternatively, they may voluntarily decide not to perceive the same augmentations.
- **Interaction and Interactivity:** Users can explore virtual content interactively with additional devices, such as Personal Interaction Panel [226].

Szalavári et al. [225] argue that AR systems can improve the viewing and manipulation of complex models when compared to traditional computer setups with a mouse and a keyboard. They also defend the idea that direct and indirect communication is easily possible with an AR system and “are probably richer than any computer-governed interaction can ever be.”

At about the same time, Billinghurst et al. [10] explored differences in collaboration between a fully virtual and an augmented reality environment through the concept of *Shared Space*. They point out several advantages of the AR environment over its fully virtual counterpart:

- Local participants can see each other's facial expressions, gestures, and body language. These possibilities increase the bandwidth of communication among participants as they provide richer channels of natural non-virtual communication.
- Participants can use familiar physical-world tools to manipulate virtual images, which increases the intuitiveness of the user interface.
- Users can refer to notes, diagrams, books, and other physical objects while viewing virtual objects.
- There is no need to model the entire environment, thus, graphics rendering requirements are considerably reduced. Augmented Reality can be used to enhance the complexity that already exists in the physical world instead of replacing it entirely.

The authors studied the impact of hiding aspects of the physical world when users are immersed in it. They observed that collaborators performed more quickly when they could see their collaborators in the physical world. This observation is consistent with results by Kiyokawa et al. [128], who found that users perceived gaze cues better in an AR setting than in VR.

These results encouraged further research on collaborative AR applications. For example, the *Collaborative Web Space* by Billinghurst and Kato [11] was a shared space in a three-dimensional web browser that enabled co-located users to browse the World Wide Web collaboratively. Web pages were displayed around the users as 2D virtual windows, and collaborators communicated around them by using voice, gaze, and gestures. By preserving the communication channels of traditional collaboration, AR technologies can transparently support natural human behavior [14].

Shortly after, Broll et al. [21] presented a collaborative AR system named the *Virtual Round Table*. The system allows multiple collaborators to see a shared virtual 3D model while they sit around a round table. The authors show that the additional perception of virtual content can maintain and even improve traditional workplace situations. Collaborative AR around a table can also be used to support complex design and planning decisions for design professionals like architects, going beyond early paper-based sketches. Extending their earlier work, Broll et al. [20] introduced a system (named *ARTHUR*) for architectural design sessions and review meetings. The system can be used with other CAD software and simulation data. Overall, it changes how designers collaborate beyond tabletops and desktop design tools.

ARTHUR [20] uses computer vision to retrieve user input through hand gestures and fingertip tracking. It also uses *placeholder objects* that support two planar and one rotational degree of freedom and *pointers* that support five degrees-of-freedom interaction tools with buttons for picking and selecting operations. But later, systems used mobile devices to enable collaborators to create virtual content in AR. For example, Huo et al. [100] used multiple smartphones to register and synchronize the reconstruction of the AR environment of the collaborators. Chao et al. [29] used smartphones as additional input for interaction, whereas Gugenheimer et al. [86] used a VR controller.

Finally, other researchers have looked at how the time and spatial proximity of collaborators affect their experience. Guo et al. [88] suggest that co-located-synchronous collaboration with mobile AR improves communication and coordination compared to remote or asynchronous collaboration. They observed more engagement, more creativity, and closer collaboration among users.

2.2.2.1 MR collaboration settings

Using a head-mounted display to show AR content is the dominant approach today. However, it can make users feel distanced from their collaborators and severely impact their perceptual cues. HCI research has explored a range of alternative technologies, such as handheld or flat panel displays [12]. Virtual content can also be generated through volumetric displays. Grossman and Balakrishnan [85] explored new interactions techniques for such devices: highlighting and textual annotations tools, object manipulation by combining 3D menus, ray casting, and a 3D cursor, and a scene-splitting technique that divides the virtual scene into multiple viewports. Interestingly, the authors noticed that users were sometimes reluctant to perform virtual navigation as they did not want to change the viewpoint of their collaborator. This observation suggests that users often verbally share their navigation intentions before they actually realize them. Volumetric displays show three-dimensional virtual content without requiring the collaborators to wear headsets. However, such devices are limited by the size and lisibility of their visualizations.

Others [208] have investigated how to enable AR collaboration with handheld tablets in room design scenarios. Their tablet user interface showed an augmented video, where the live stream video captured by the tablet's camera was combined with virtual content. The authors ran a user study to explore the use of this system. Although collaborators could observe their partners' manipulation actions from their own screen, they often looked at each other's screens. Likewise, the authors observed that during the design task, some users switched from using their own tablet to sharing a single tablet with their collaborator.

Various ways of experiencing AR have been developed, where collaborators are not tied to use identical devices or the same interac-

tion configurations. Thus, we can distinguish between *symmetric* and *asymmetric* collaboration setups. Previous HCI research has explored various forms of asymmetric collaboration. Grandi et al. [83] investigated asymmetric collaboration between a user wearing a VR headset and a user with an AR tablet. Both interacted with the same virtual content but with different devices and tools. Grandi et al. found that the VR-AR asymmetric configuration led to a better performance than an AR-AR symmetric configuration but also to a slightly worse performance than a VR-VR symmetric configuration. They observed that cooperation levels were overall unaffected by the symmetry or asymmetry of the collaboration setup. Yet, they found that participants engaged in slightly less mutual assistance in the asymmetric condition.

Mackamul et al. [155] studied the effect of using handheld AR devices and projection-based AR on a memory game of match pairs. They found that participants performed more card flips but finished the game quicker when using projection. However, interaction with handheld devices was described as easier, more intuitive, and more familiar.

Gugenheimer et al. [86] studied how to overcome the communication barriers that appear when a virtual space is not shared among collocated users. They introduced *ShareVR*, a system that makes the virtual world of a VR user accessible to a non-VR user via floor projection and a mobile display situated on an HTC Vive controller. The authors found that their system improved enjoyment, social engagement, and presence compared to a more traditional system, where the external user views and interacts with the virtual scene through a Gamepad and a TV.

Other systems let non-VR collaborators interact with the virtual scene through a simple mobile display. For example, *TransceiVR* [239] uses a tablet to view and interact with the virtual world of the user who wears the VR headset. In contrast, *ShARe* [105] enables collaboration between AR and non-AR users using projection. But collaborators can also use different devices to collaborate seamlessly in AR and VR over the same virtual content in the same physical space [194]. Finally, Wells and Houben [252] showed that collaboration with mobile AR interfaces could lead to high physical and cognitive load. In particular, they observed that participants had to frequently switch their attention between the AR model on the mobile device and their collaborators when they worked together face to face. Their results suggest that the quality of collaboration may decrease as the complexity of the virtual model increases.

2.2.2.2 *Workspace perception*

The location of the workspace relative to the positions of the collaborating users can also impact the collaboration experience and performance. Kiyokawa et al. [126] compared three different config-

urations of workspace locations: among users, on top of a table that is placed between the users, or on a wall on the side of users. They found that under the first configuration, communication became "more natural, social and easier" and "subjects made more initiatory body motions, utterances and laughter" [126].

But collaborators can change the perspective in which they view their workspace, for example, by rotating it or by moving around [33]. Billingham et al. [12] report that when users can move around a table with virtual objects, their interaction with these objects is easier than when the two users are side-by-side in front of a projection of the virtual content. The task space can also be mobile. For example, Regenbrecht and Wagner [188] introduced the "Cake platter" technique, which uses a tangible turntable plate-shaped device on top of which 3D virtual content is displayed (see Figure 2.10).

Another approach consists in representing some area of the workspace through virtual proxies that extend the interaction capabilities of users [261].

Chastine et al. [33] observe that deictic gestures are important in collaborative AR environments for establishing a common ground. For example, they observed that the majority of participants preferred using their hands rather than a virtual arrow to make references. Chastine et al. highlight two key points for cooperative collaboration in AR:

1. Pointing benefits from the use of natural hand gestures; and
2. Shared viewpoints are effective for communication when establishing common references.



Figure 2.10: Four users looking at a common 3D model placed on a rotating "cake platter". Figure from Regenbrecht et al. [188].

These points are consistent with earlier observations by Tang et al. [230], who state that hand gestures can help collaborators generate and interpret reference cues. Chastine et al. [32] confirm these points by studying the role of pointers in AR, and suggest a need for shared viewpoints.

2.2.3 Remote collaboration

Real-time collaboration is a challenge when users work remotely and, consequently, do not share the same physical environment. Such situations became commonplace during the COVID-19 pandemic [271]. Many design and research teams have found themselves working remotely, relying on video-communication software to collaborate together [258]. Some experts predict that such situations are not temporary and will largely persist after the pandemic [15].

AR supports remote collaboration tools that go “*beyond being there*” [97], and collaborators can use them to work on augmented content that is either onsite or remote [262] (see Figure 2.11). Back in 1999, Billinghurst and Kato [11] introduced *WearCom*: a portable MR interface that supports remote collaboration between a single wearable computer user and multiple desktop users. This system provides a shared communication space among collaborators. Remote collaborators are represented as audio-enabled avatars placed on a cylinder surrounding the user with the AR headset. Its goal is to make a seamless connection with the physical world and improve realism compared to desktop-based collaboration environments.

However, the above approach brings two fundamental questions:

1. When users collaborate remotely, how can we preserve (or further enhance) the communication references that naturally exist in their space while taking advantage of their communication gestures?
2. What is the best representation for remote collaborators and their workspace?

In the following paragraphs, I discuss previous work that deals with these questions.

2.2.3.1 Sharing references through annotation, pointing tools, and gestures

For remote collaboration with Augmented Reality, Chastine et al. [33] highlight the need for the following functionality:

- the ability to point to objects or locations in the scene;
- reference points embedded within the environment that ease communication among users;



Figure 2.11: Two AR collaborators working respectively from a) on-site and b) remote. Figure from Zhang et al. [262].

- effective use of video sharing to provide additional communication channels for selection and reference.

The authors observe that while references made with the users' hands are important for co-located collaboration around physical objects, when working with virtual objects, or in remote collaboration (in particular, when hand-based references are not available), virtual pointers can act as alternative reference tools that ease communication among users.

Likewise, pointing tools can augment the physical environment and facilitate remote collaboration. Kervegant et al. [121] use a *Distant Assist Cursor* to let remote desktop users point to a position in the space of the local user with a projected dotted light. The desktop user perceives the local workspace through a camera. According to Kim et al. [123], using virtual pointers and annotations in video-mediated communication can help collaborators feel more connected and closer together and better understand their partner.

However, a simple cursor-pointing tool is often insufficient for remote collaboration on physical tasks. Fussel et al. [68] consider a *cursor* and a *pen-based drawing tool* instead. Such pen-based drawing tools "can lead to communication and performance virtually identical to that found in side-by-side collaborations.". Likewise, Seungwon et al. [122] suggest that shared drawing annotations can improve user performance compared to only using pointers.

Annotations can provide additional layers of information in a remote workspace. There are various solutions on how to integrate annotations into a remote workspace. For video-mediated communication, Barakonyi et al. [9] showed how to integrate 3D virtual content into a video stream. This approach can then be used to share annotations over the video stream, for example, between a desktop computer user and a remote tablet user [74]. Speicher et al. [218] proposed instead a

framework for remote mobile collaboration based on an augmented 360° video stream that lets users annotate its content.

Several studies have shown the added value of annotation tools in remote collaboration. Velamkayala et al. [242] studied a collaborative navigation task between a user with a tablet and a wearer of an AR headset through videoconferencing software. They found that when tablet users could draw annotations in the field of view of the AR user, collaboration appeared to be slower but more precise than when using videoconferencing with smartphones only. Besides, Gauglitz et al. [73] found that spatially anchored virtual annotation of a live video stream helps users become faster and more confident about their task performance than non-anchored or non-annotated video.

MR environments can also support remote communication through gestures. For example, in *HandsInAir* [98], an AR user can make gestures on a remote workspace retrieved via a camera. Remote AR collaborators then perceive these gestures as an augmentation that is overlaid to their workspace. A different system by Feick et al. [62] is instead dedicated to remote collaborations between a novice AR and an expert VR user, with a focus on object-oriented tasks. The system relies on proxies that allow users to perceive the view of their collaborators. But it also captures the hands of the remote experts while they point to objects with their hands.

Finally, other research has studied the use of annotations in conjunction with hand gestures. Teo et al. [236] introduced a system for remote collaboration between a wearer of a VR headset and a wearer of an AR headset, on top of whom a 360° camera is placed. The system supports non-verbal communication with gestures shown through 3D hand models for the VR user and a video stream for the AR user. A preliminary user evaluation showed that visual annotation cues (consisting of ray-pointing actions and drawing annotations) helped AR users perform tasks faster and with less error. However, no clear evidence of help was found for VR users. Clay et al. [37], in turn, studied the use of gestures and annotations (in the form of virtual arrows) when a local worker wore an AR headset with a frontal camera and a remote expert perceived its video on a desktop computer. They observed that annotations dominated the collaborators' actions.

2.2.3.2 *Perceiving a remote workspace through video*

The HCI literature has long examined the effect of different views on video-mediated communication, especially in the context of physical tasks that involve spatial object manipulation and construction. Back in the 90s, Kuzuoka [133] investigates spatial workspace collaboration through *SharedView*, a video communication system. Kuzuoka's study requires a remote expert to explain a 3D task to a local worker in a machining center and shows that the viewpoint of the video can affect the efficiency of communication. *GestureCam* [134] goes a bit

further by giving the ability to the remote operator to move a local camera to perceive the workspace, but also point with a laser or provide additional gestures on top of the video in direction to the local workers.

Gaver et al. [75] study the use of five camera views for a remote-collaboration design task. Their task requires a participant in a local office to arrange the furniture in a dollhouse in collaboration with a remote partner. Results show that participants largely preferred task-centered views to face-to-face communication. The authors also observe that view-switching can be problematic. In particular, multiple views can interfere with establishing a common frame of reference, introduce discontinuities, and impede coordination. Ten years later, Fussel et al. [67] compare two remote-view configurations: (i) a head-mounted camera with eye tracking and (ii) a scene camera placed at the back of the worker, providing a wider but fixed view of the working environment. The scene camera is shown to be preferable and improves communication efficiency, while the head-camera view does not add any benefit compared to an audio-only condition. Similarly, adding a second head-camera view to a scene-camera view seems to deteriorate rather than improve collaboration performance. A more recent study [223] in the context of AR video collaboration for 3D guidance tasks also shows that a third-person view results in better task performance and higher user satisfaction than a first-person view. Moreover, camera control and the system's responsiveness are important factors when using an augmented video stream [254]. Rae et al. [184] observe viewpoint control via robotic telepresence systems. Despite a greater sense of presence than traditional video-conferencing tools, they show that such systems also increase the remote user's cognitive load. Domova et al. [48] also explored the ability to make snapshots in a video-mediated communication regarding a 3D workspace to maintain a viewpoint while the camera is looking elsewhere.

However, other studies show the advantages of combining multiple alternative views. For example, Schafer and Bowman [201] study a virtual furniture arrangement task and observe that the availability of two alternative representations (virtual 3D and floor plan) *“enabled the users to investigate different aspects of the space”*. Ranjan et al. [186] found that remote users complete complex Lego construction tasks faster with an automatic pan-tilt-zoom camera than with a static camera. Giusti et al. [80] investigate how a local user and a remote expert configure a mobile phone and a tablet to repair a Lego model or replace a punctured bike tube. When both a phone and a tablet were available, local users tended to fix the tablet camera view to show an overview of their workspace and sometimes their faces. At the same time, they used the mobile phone camera when they needed to zoom in on specific parts to show details. Lanir et al. [137] investigate user performance and behavior with respect to who has the camera

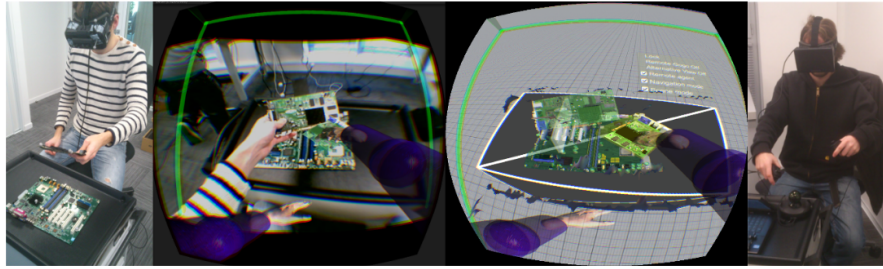


Figure 2.12: An agent (left) perceives virtual replicas of the arms of a remote expert (right) for a motherboard assembly task. Figure from Le Chenechal et al. [36].

control (the local vs. the remote user). They conclude that the outcome depends on the situation and the task at hand. Overall, results are far from conclusive but seem to suggest that the most suitable strategy is to give users control over alternative views, each adapted to a different type of task.

2.2.3.3 Representing remote collaborators

Several MR systems rely on video to help remote collaborators view each other or, at least, view parts of their body as those interact with the objects in a scene. For example, Robert et al. [191] studied collaboration between a user with an AR headset and a user with a tablet who tries to provide remote instructions. In their system, video from the AR headset is streamed to the tablet, also showing the hands of the remote user. Similarly, *SharedSphere* [143] lets VR users get immersed in the live stream of a 360 camera worn on the head of their AR collaborator. The VR users view their collaborators' hands from the video, while the AR users perceive their collaborators' hands via a virtual 3D representation.

Virtual representations of remote users are common in several other systems. Chenechal et al. [36] show the virtual arms of the remote VR user in the field of view of a local AR user (see Figure 2.12). Piumsomboon et al. [181] represent remote users with a virtual 3D head and a virtual hands model. Bai et al. [8] apply a similar representation but also add a virtual line to represent the eye gaze of users. Some systems construct the mesh of a user's avatar from video frames [61]. This representation encourages non-verbal communication as it provides rich information about the full body of users. Other approaches [234, 238] build instead a point-cloud representation of the collaborators' body.

Although such representations of remote collaborators are meant to be useful, they may not always be relevant. Osmers et al. [171] found that the presence of an avatar has only an effect if the task or the setting creates a need for them.

2.2.4 Conclusion

According to Gutwin and Greenberg [89], “workspace awareness is useful for making collaborative interaction more efficient, less effortful, and less error-prone”. When collaborators are physically distant, preserving awareness is more challenging. Still, there are many possible ways of supporting awareness. For example, pointing or annotation techniques can offer spatial references for collaboration. There are also techniques that communicate hand gestures in the form of augmented video, mesh, or point-clouds. Unfortunately, such techniques have their own limitations, as they cannot convey all information cues that are naturally available when people interact within the same physical space.

I also reviewed techniques that rely on video-mediated communication, using one or more physical cameras. A main limitation of such techniques is the fact that the views of remote users are constrained by the viewpoint of these cameras. For example, a head camera is only controlled by its wearer. We need to better understand how to best make use of them.

Furthermore, I discussed the role of territoriality and examined strategies for managing collaborators’ occupation of space. While several studies have explored collaborators’ behavior regarding private and public space, there is very little related work on collaboration in AR environments, where the workspace is 3D and virtual content appears in a shared physical space.

2.3 BEYOND PHYSICS

Physical laws limit our potential for action in the physical world. They are unbreakable and apply to us as well as our surroundings. For example, gravity makes us tied to the ground, and walls won’t let us cross them. Likewise, our perception of the world is limited by physics, so we see only what is directly reachable by our eyes. And as distances become larger, our direct communication channels through voice, gaze, and gesture disappear.

To some extent, interaction design has transferred physical behaviors to the virtual world [257]. Previous research has even tried to apply physics to WIMP interfaces on the desktop [2]. In the context of collaborative systems, Randall B. Smith argues in favor of a “*What You See Is What I Think You See*” [212] principle, which requires that collaborators interact as in the physical world.

There are many reasons why preserving the metaphors of the physical world might be desirable. First, people are familiar with physics and can predict its outcomes. Thus, breaking their laws in a virtual 3D space could bewilder users. Adapting to a new behavior requires learning, which means additional time and effort. Also, going from

reality to virtuality, or inversely, from virtuality to reality, requires a certain consistency such that users can easily understand how to adapt their actions. As reality is constrained by its physics, by transferring that physics to the virtual world, we can achieve fluid transitions that leverage the existing skills of users [194].

However, virtual worlds are completely artificial and do not need to be subject to physics. So it is possible to bypass the constraints of the physical world and explore new design possibilities, e.g., create novel VR interactions that extend the laws of physics [257] [88]. This can also be true for AR environments, where virtual spaces are embedded in the physical world. As Michelis et al. [45] point out:

"Augmented places are not virtual copies of a physical place: their spatial arrangements are not possible in the physical reality, they are the genuine outcome of the design of a new form of space, with qualities like plasticity and flexibility quite absent from the physical space where we live" [45].

Thus, AR can also support non-physical interaction behaviors.

In the following paragraphs, I will discuss ways in which virtual behavior goes beyond our expectations when we interact with the physical world. I will present interaction techniques and virtual representations that challenge the physical laws. I will especially focus on techniques that allow remote collaborators to deal with the distance that separates them and work together. Finally, I will take a closer look at how users can navigate among multiple-view representations to overcome the physical constraints posed by their viewpoints.

2.3.1 From physical to virtual spaces and objects

A virtual space in which interaction takes place can have properties different than its physical counterpart.

As an example, Thanyadit et al. [237] use the *mirror metaphor* to design a virtual surface that mirrors users' hands and reproduces their movements. The users can then manipulate the objects in front of them with their physical hands and the objects behind the virtual mirror with their mirrored hand replicas. This interaction is then used to organize files, where the space in front of the user is a working space, and the space of the hands' replica is a storage space.

Also, interaction in a virtual space may involve "unusual" objects, where virtual content can take interactive representations that are different from how it would normally appear in the physical world. For example, Sihwa Park [174] represents a recorded audio signal as a three-dimensional shape in AR. Users can then interact with this shape to replay the recorded signal or change its characteristics. Or, as in GhostAR [26], movement can be represented as a transparent volume.

Representing such abstract non-materialized concepts as virtual objects provides new ways of interacting with them.

The decorrelation between virtual spaces and objects in the physical world may change how collaboration takes place. For example, abstract objects can be used to populate the virtual space of users who interact through an AR tablet. Used as *Shared Virtual Landmarks* [165], they can positively influence communication behavior. They may reduce the occurrence of ambiguous deictic expressions and avoid potential conflict situations. Similarly, in a co-located collaborative mixed-reality task for object identification and positioning, Muller et al. [164] observe that adding virtual objects as spatial cues positively influences collaborators' communication behavior, decreases user task load, and improves user experience.

2.3.1.1 *The integrity of virtual objects*

Physical objects need to be rebuilt to regain their original state after modification. In contrast, the geometry of virtual objects can be altered without affecting their original shape. We can thus design techniques that can break or deform a wall to look inside hidden objects. For example, Pindat et al. [175] introduce a visualization technique that enables users to “dig” a temporary hole in a virtual 3D model to observe its interior. McGuffin et al. [160] take another approach, where they deform the model to visualize its content better. They consider various deformations, such as: defining a volume around an area of interest and opening this volume like a peel (Peeler Tool), rotating this area around a given axis (Leafer Tool), expanding surfaces from a given point (Sphere expander tool), remove content inside a box (box spreader tool) or slice a volume using a v-shaped virtual knife (Hinge Spreader Tool).

Another approach consists in defining a volume and selecting which part of a 3D model is rendered inside and outside. Viega et al. [244] first demonstrated this technique using Magic Lenses. Magic lenses can be flat or volumetric. Looser et al. [150] further explored this technique for AR interactions, such as *Magnification*, *Object Selection and Manipulation*, or *Information Filtering*. Such interactions make it possible to respectively examine distant objects closely or close objects in greater detail, select and manipulate objects visible in the lens, and filter the information shown through the lens. Lenses have been studied for object selection. The results show that while lenses are not especially accurate, they require less hand and head movement with lower head movement speed over direct touch or ray-casting [151]. The shape and tangibility of such lenses have also been explored with *Flexible AR Tangible Magic Lens* [152]. This interaction is based on a foldable physical surface that the user can move around using two physical handles (see Figure 2.13).

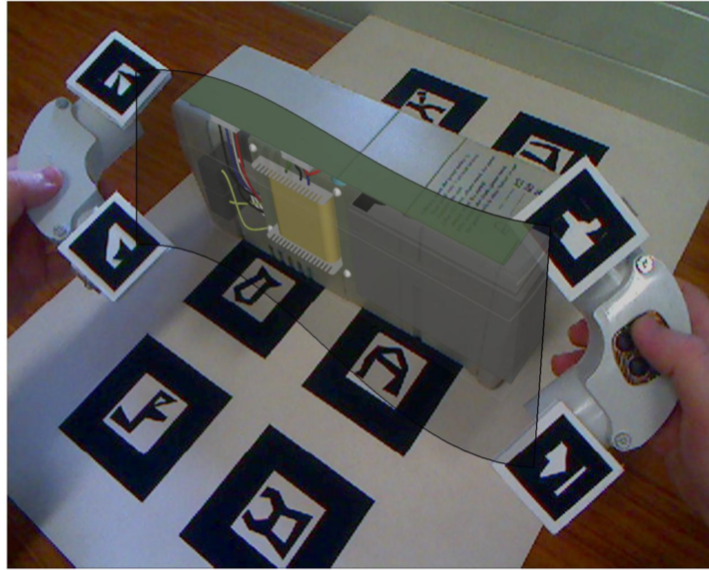


Figure 2.13: A flexible and tangible Magic Lens used to investigate the internal components of an electronic appliance. Figure from Looser et al. [152].

In Mixed Reality, a physical object can have a digital counterpart. These virtual objects may, for example, be used to generate instructions for guidance by visualizing the final placement of a physical object [168]. They can also show certain aspects of a physical object without hindering its manipulation. In a remote Mixed Reality scenario, a user can then manipulate a physical object while their collaborator manipulates a virtual copy. This helps perceive different perspectives of this object among collaborators and facilitates deictic references [62]. Likewise, a part of a physical object can be duplicated to provide an additional alternative view and help physical operations in occluded areas, like connecting a plug behind a television [147]. Virtual objects can also be duplicated to avoid manipulation conflicts in a collaborative scenario. Xia et al. [257] use, for example, *parallel objects*, which are copies of virtual objects created when collaborators disagree on where this object should be placed.

2.3.1.2 Transparency

The colors and transparency of physical objects are defined by their materials. As physical matter is defined by its physical properties, it can be hard to modify. In contrast, virtual objects are not made of such matter. Therefore, their representation can easily be controlled and changed through time.

Partial transparency can temporarily unveil occluded areas. Elmqvist et al. [57] show that such *dynamic transparency* can lead to faster performance and more precision in 3D navigation compared to no transparency. Also, partial transparency improves the perception of

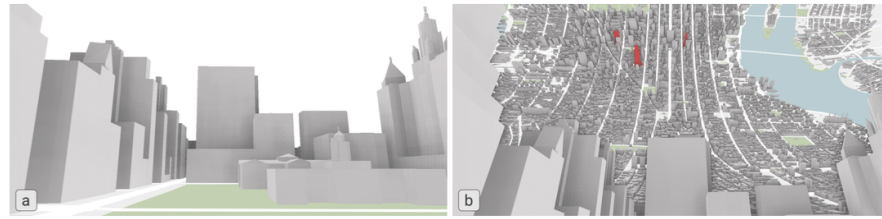


Figure 2.14: Geometrical deformation of a virtual city model (a) to extend the visibility of occluded streets and buildings (b). Figure from Chen et al. [35].

important occluded features while still showing elements from the occluding part of the model [112]. Transparency can be simulated in AR to provide a *X-Ray* effect and help spatial navigation [47]. Similarly, Klemen et al. [147] found that seeing through an object using AR helps perform operations in occluded areas.

Transparency can also help users perceive multiple spaces at once when it is applied to the whole scene. Schjerlund22 et al. [202] used transparency to represent multiple virtual places simultaneously. By overlapping virtual spaces and making them transparent, those become visible at the same time from a first-person perspective. The authors found that this technique reduces completion time and the physical motion exerted by the user's head in an object collection task.

2.3.1.3 Workspace deformations

There also exist techniques that deform the visual space to help users get a wider view of their workspace (either personal or shared). Fraser et al. [65] used *Peripheral Lenses* [192] to extend the peripheral vision of users in a desktop Collaborative Virtual Environment. Such lenses may increase awareness about the workspace. However, their peripheral distortion may hinder interaction and may fail to represent collaborators' actions in a precise way.

Another approach consists in deforming the virtual content itself. *UrbanRama* [35] includes an interaction technique that folds the space in front of a VR user (see Figure 2.14). This spatial deformation lets users perceive occluded elements and visually explore urban landscapes in VR. By mixing local and global perspectives of city 3D models, Chen et al. [35] show that the technique does not hamper the user's perception of distance orientation and does not degrade the navigation experience. *Ballon probe* [56] inflates the virtual space at a given point in space and avoids occlusions by taking away overlapping objects. Unfortunately, such space deformations alter the relative positioning of objects and may disorient users.

The virtual representation of a user's body can also be altered. In particular, the position and orientation of virtual collaborators can be

manipulated to be slightly different than the physical ones. Previous work [159, 217] suggests that controlling this difference according to the pointed location and the viewpoints of collaborators can help increase the accuracy of deictic gestures.

Deformations can be extended to the full virtual workspace. For example, Pisani et al. [177] experimented with hyperbolic virtual spaces that do not respect Euclidean geometry. They suggest that such spaces may feel more intuitive for navigating branching structures. Singh and Balakrishnan [210] went even further by exploring the visualization of 3D scenes with non-linear projections.

2.3.1.4 Collaboration at different scales

Although changing the size of the physical world is impossible, virtual environments are not bounded to this constraint. Thus, we can build VR collaborative environments whose users can experience different scale factors [64]. In particular, VR users can change the size of their environment to get a better view of a 3D model. Previous work has introduced various VR and AR systems that play with the scale in imaginative ways. Roo and Hachet [195] switch from AR to VR to scale virtual objects in mid-air for better inspection and closer interaction. In a collaborative face-to-face AR scenario, Kiyokawa et al. [127] enabled users to seamlessly switch from AR to VR to scale the workspace and change perspective as they work around a shared 3D model. Piumsomboon et al. [181] extend this concept further by supporting awareness cues about users' hand manipulations and their gaze. They define a *God Mode*, where the VR user appears at a bigger scale than the AR user, and a *Miniature Mode*, where the VR user becomes smaller than the AR user. Such configurations had been initially explored more than ten years earlier by Stafford et al. [220]. Their *God-like interaction* helped tabletop users to provide deictic instructions to their AR collaborators at a different scale (see Figure 2.15a). The AR users perceived their collaborator's hand as a giant hand that moved over their head (see Figure 2.15b).

In *SnowDome* [178], which is a more recent system, VR users can interactively scale their environment by going into or out of a pre-defined area of their workspace. Or they can use it as a miniature mock-up [41].

In a virtual environment, collaborators cannot simply change the scale of the objects. They can also interact with and scale the representation of their peers. For example, Piumsomboon et al. [179] use avatars to represent VR users in the virtual space of their AR partners. When the VR users are out of sight of their collaborator, their avatar is automatically scaled down and moved to a visible location. Piumsomboon et al. report that this scaling technique improves the performance of object placement and enhances social presence. Xia et al. [257] make avatars first-class interaction objects (*Avatar Objects*)

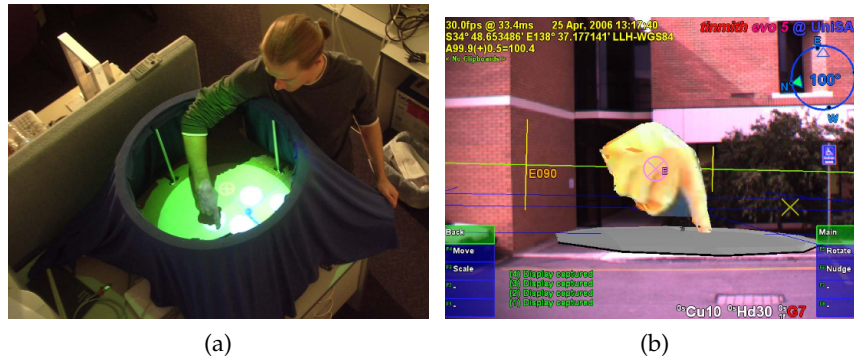


Figure 2.15: (a) An indoor user pointing at a location on a tabletop surface, which contains a virtual representation of the outdoor world. (b) Outdoor AR view showing the indoor user's hand appearing from the sky and pointing toward a location in the distance. Figure from Stafford et al. [220].

that allow VR users to manipulate their collaborators like other virtual objects, e.g., they can grab them, make them smaller, and bring them closer to show them their edits. Piumsomboon et al. [180] apply a similar approach to AR: a VR user (Miniature) is immersed in a 360 degrees live video stream of a physical-world scene captured by a camera that is controlled by a remote AR user (Giant). The camera can be worn on the shoulder or the head of the Giant or be manipulated directly with their hands.

2.3.2 Reconstructing a remote MR workspace

To perform a collaborative task in an AR environment, the users have to share information from both the virtual and the physical world. For co-located collaboration, this sharing is implicit because collaborators perceive the same physical and virtual workspace. They can therefore rely on traditional communication channels, such as voice, gaze, or gesture. However, when collaborators are not physically present in the same workspace, communication needs to be performed remotely. Likewise, when remote users cannot be physically present, they need to have access to synchronized representations of this workspace. Previous work explored different approaches to address this situation, either through video or by reconstructing the indoor environment of the local user [114]. I have discussed video-mediated communication systems in previous sections (see Section 2.1.2.3). Here, I will focus on 3D reconstruction techniques and highlight some of their limits.

2.3.2.1 Mesh reconstruction

A common solution for supporting remote collaboration is to use virtual replicas of physical objects, as in the work of Oda et al. [168]

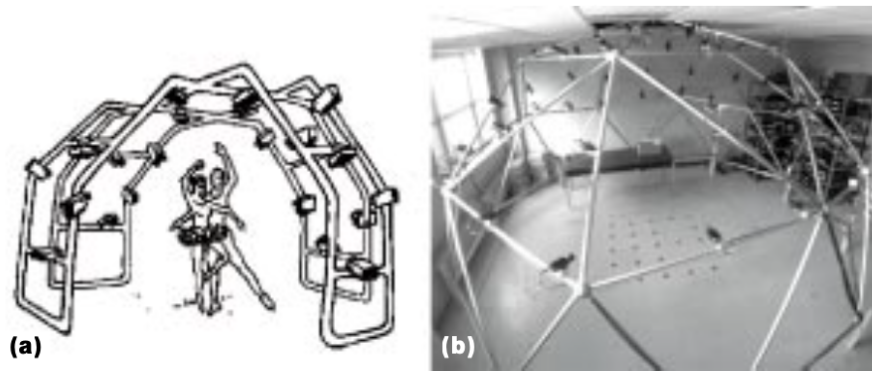


Figure 2.16: The Virtualized Reality studio: conceptual (a) and technical implementation (b). Figure from Kanade et al. [113].

where a remote expert wearing a VR headset gives instructions to a local worker with an AR headset. Unfortunately, such virtual replicas have to be created in advance and provide partial only information about the physical environment of the local AR user.

It is possible to reconstruct a physical environment in a virtual world by using a set of static cameras. Fuchs et al. [66] propose a "sea-of-camera" for a 3D teleconferencing scenario. These cameras create a polygonal model of the workspace that is visible to all collaborators. In the same direction, the *Virtualized Reality* reconstructs a physical scene in the virtual world using a cluster of cameras placed around a dome. An early example of such remote workspace reconstruction was the *Virtualized Reality studio* [113], based on 51 cameras placed on a geodesic dome (see Figure 2.16).

More recently, Gauglitz et al. [74] used a simultaneous localization and mapping (SLAM) system with a tablet or smartphone camera to project three-dimensional annotations. Such techniques can be combined with AR headsets [264] and can support a range of scenarios, such as helping remote users configure IoT devices [154]. Their key limitation is that they consider spatial information as "seen" by the device and thus provide partial only reconstruction. Other research on remote mobile assistance investigated how to track a 3D space through light fields [163]. However, such systems lack depth information and occlusion management can be problematic.

Finally, workspace reconstruction can rely on fusion techniques [104]. In *BeThere* [213], for example, a local user pre-captures the 3D geometry of the workspace with a mobile depth camera to share it with a remote user. Piumsomboon et al. [181] followed a similar reconstruction approach to share a physical space with a remote VR user. Adcock et al. [1] used instead multiple depth cameras to implement a real-time volume fusion of the physical space. Their system constructs a colored mesh that can be updated through time to track dynamic objects. Unfortunately, this fusion of a dynamic space is not immediate. Despite promising research on dynamic mesh reconstruction for the

human body [50] [209], such reconstruction techniques cannot provide yet a dynamic reconstruction of a full workspace.

2.3.2.2 *Point-cloud representations*

An alternative solution for representing dynamic physical-world 3D scenes is through point clouds. In the context of remote collaboration, point clouds have been used to communicate the physical environment of a local user, part of the collaborators' body, or specific objects in a physical workspace. For example, Gao et al. [72] visualize with point clouds the hands of remote users. Bai et al. [8] use them to communicate the physical space of an AR user to a remote collaborator who wears a VR headset [8]. To do so, they install multiple depth cameras in the workspace of the AR user. A limitation of their system is that it can only display low-resolution 3D panoramas and simplistic avatar representations of the local user.

Gao et al. [71] use a single smartphone with a depth sensor to reconstruct a physical scene as a static point cloud. The mobile user captures the local scene from multiple viewpoints to generate a colored point-cloud representation. A remote VR user can then interact with this representation to inspect the workspace of the local user and provide guidance.

2.3.2.3 *Mixed representations*

Jones et al. [110] observe that the reduced quality of a full 3D reconstruction can distort collaborators' expressiveness and make them experience an "*uncanny valley of XR [extended reality] telepresence*". The authors also report that "*the more immersive an XR Telepresence system is, the more amplified technical issues, such as latency, video quality, and control become*" [110]. A complementary strategy consists thus in combining multiple representations to counterbalance their limitations.

Kumaravel et al. [238] study two representations that communicate the virtual and physical workspace of a local user: (i) a 2D video stream and (ii) an *hologlyph*, a 3D representation of spatial data captured by depth cameras and rendered as a point cloud. Feick et al. [62] combine, instead, two parallel views for a remote expert user: (i) a video feed showing the other user manipulating a physical object, and (ii) a 3D scene that allows the expert to gesture over a virtual proxy of the object. In other mixed-reality systems, remote collaborators can switch between a 360° panorama video and a 3D reconstructed scene [235]. Gao et al. [70] used a static 3D reconstruction as well as 2D or 360° video to support a remote collaboration between a VR and an AR collaborator. They found that the VR collaborators preferred the 3D static overview over the video to search for targets and learn the physical layout of a remote workspace. The 360° live video view was preferred over the 2D first-person view. The main reason was that the

former provided more independence, thus, users could better control the task process and check their collaborator's actions.

2.3.3 View transition techniques

VR worlds impose few constraints on the way users can navigate in 3D space. This navigation freedom enables them to inspect virtual content from arbitrary viewpoints [1] [113]. It also helps them comprehend movement in space, as actions can be replayed from different perspectives [43] [209].

Viewpoints can also be controlled by creating portals, as in Photoportals [132], or by teleporting collaborators, as in Spacetime [257]. These view transition techniques, however, support navigation within virtual environments only – they do not consider the collaborators' physical space. But other authors have tried to address this gap. For instance, Tatzgern et al [232] studied how to seamlessly transition between AR and VR views, taking into consideration both the virtual and the physical world. In other cases, the users may need to be present in the same physical space but interact with different virtual or augmented environments. For such scenarios, Slice of Light [248] uses physical proximity to enable users to view and navigate within the virtual space of their collaborators.

JackIn Space [129] propose a spatial navigation technique for collaboration that simulates a virtual out-of-body experience. A VR user can seamlessly transition from a free third-person view of the workspace to the first-person views of their collaborators (see Figure 2.17). By doing so, they switch from the view of fish-eye cameras worn by collaborators to a virtual camera view showing a point cloud reconstruction of the workspace. But in contrast with VR, AR collaboration is constrained by the position and coordination of the available physical cameras. With the *MagicBook*[13], AR users can switch among augmented content by turning the pages of a physical book but users have to move around to change their perspectives. Therefore, they are constrained by their physical abilities.

Using multiple cameras can provide alternative viewpoints. Transitioning among these views can help in teleoperation, such as controlling a robot. Praveena et al. [183] suggest that switching among multiple view-aligned control frames may allow users to work in a local frame of reference without maintaining global spatial awareness. In Augmented Reality, Rasmussen and Huang [187] show previews from multiple cameras to a remote desktop user. They can then switch among them to perceive a remote workspace from different viewpoints.

View changes in multiple video streams can be done manually or automatically. In the first case, viewers have autonomy and freedom, but this induces a higher task load for them. In the second case,

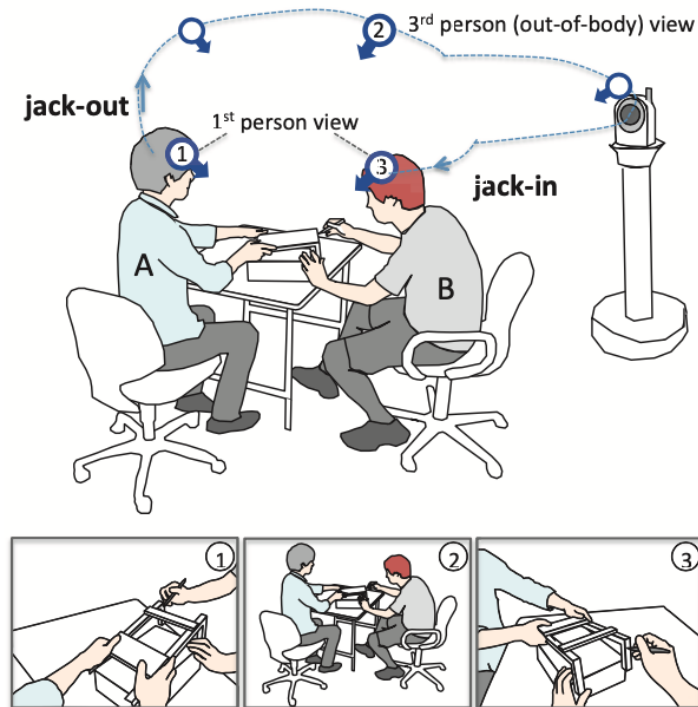


Figure 2.17: The JackIn Space telepresence concept: (1) A remote user perceives A's first-person view (view 1). (2) The remote user's viewpoint smoothly moves out from the viewpoint of A to become a third-person perspective (view 2). (3) The user can freely change his/her view position, and alternatively, see from external cameras. (4) The user's viewpoint moves to user B's viewpoint (view 3). Figure from Komiyama et al. [129].

viewers experience less control but can concentrate on the content rather than following an object of interest [241]. Tait et al. [227] found that view independence can increase the speed and quality of collaboration when instructions are given from a remote desktop user. However, view independence does not necessarily improve the accuracy of the performed task.

Finally, changing views can help AR users keep track of time. e.g., by quickly switching among snapshots of their past views [222].

2.3.4 Conclusion

I discussed how interaction with virtual worlds extends beyond what is possible in the physical world. A virtual environment can support imaginative representations and is subject to a range of transformations that can change its geometry, scale, and transparency. These transformations can be used to perceive occluded areas or layered 3D models better. In AR environments, however, such techniques are not easy to implement. While a virtual augmentation can be trans-

formed without considering the physical laws, the real objects cannot be deconstructed, scaled, or made translucent.

A possible solution is to use 3D reconstruction techniques that capture the geometry of the remote physical space and its objects. 3D reconstruction can enhance collaboration by offering more advanced interactions with the environment of remote collaborators. It can also support new representations that extend the physical laws, e.g., through transformations that avoid occlusions or better manage the available space. Unfortunately, 3D reconstruction techniques require complex hardware setups and bandwidth-consuming systems. Furthermore, the collaboration efficiency in reconstructed environments depends on the quality of the actual reconstruction, which is often not realistic.

An alternative solution is to combine multiple views or representations. For example, a MR system could let users see augmented video streams from multiple cameras in the distant workspace. Such videos would then be combined with virtual scenes. The downside of this approach is that switching among representations breaks the continuity of users' perception and may increase the effort required to navigate the scene. Thus, the challenge is to provide the right techniques that deal with this problem.

2.4 POSITIONING MY WORK

In the rest of my dissertation, I study two distinct situations of collaborative design. First, I focus on co-located collaboration, where designers meet at the same physical space. In this case, they work in close proximity with each other and use natural speech, gaze, and gestures to communicate directly. But in an AR environment, the virtual worlds the collaborators see and interact with do not have to be the same. The "What You See Is What I Think You See" principle does not necessarily hold. I am especially interested in solutions that assist the early design phases, where designers need to generate a large number of ideas, compare them, and then discuss them together.

I borrow ideas from space management techniques developed in other contexts (e.g., interaction around tabletops [206]), which enable collaborators to switch between a personal and a shared workspace. I also draw inspiration from VR frameworks like Spacetime [257] that introduce concepts of collaborative interaction objects that break the limits of physical laws. But when users wear an AR headset, their personal spaces can be hidden to others, so the key challenge is how to preserve awareness about the tasks of co-located partners, help them effectively switch among their virtual spaces and coordinate their actions.

Second, I explore how AR can assist remote collaboration design tasks. A few months after my thesis started, the COVID-19 outbreak

forced millions of people to work at home. Video became the dominant medium of communication, and I personally had to collaborate remotely with my thesis advisors. In our remote meetings, I had to explain my design solutions, but this was not always easy as they did not have access to AR equipment. This situation further motivated me to look into this problem. Unfortunately, video does not convey all communication cues available in face-to-face collaboration. And the situation can be more problematic when remote users need to collaborate around a 3D design task, move in space, annotate or interact with a 3D model.

In the previous sections, I identified existing approaches to this problem, but I also discussed their limitations. Some approaches rely on 3D reconstruction techniques but require a complex hardware set-up to fully capture a 3D space. In addition, their degraded reconstruction quality can hinder collaboration efficiency by letting users experience an “uncanny valley of telepresence” [110]. Thus, I focus on more practical solutions based on the use of multiple virtual or augmented views that complement each other. I also reviewed the results of several studies [67, 75, 223] that compared the effectiveness of alternative camera views. However, most of these studies have looked into traditional video representations without augmentation. Besides, to a great extent, their conclusions have been contradictory and non-conclusive. My goal is to better understand the trade-offs of complementary views and representations (either virtual or augmented) and investigate techniques that could seamlessly link them together.

CO-LOCATED COLLABORATION WITH PARALLEL VERSIONS

AR technologies have direct applications to traditional design sketching and modeling. The embedded virtual space offers a three-dimensional canvas for generating and modifying 3D objects. Because AR users also perceive their physical surroundings, they can create and modify virtual 3D content by using physical landmarks, materials, and tools as context [135] or as guides [245]. Making use of the physical environment is particularly useful, for example, in architectural design [20], when designing design of a living room [208], or in manufacturing [251].

In design practice, designers usually start by sketching ideas [82]. This ideation phase is usually conducted in groups, which allows designers to take advantage of their different skills and expertise. Technology can support this collaborative process through software tools that help the designers to develop, compare, and confront ideas. For example, systems like Teamstorm [92] support the collaborative creation of parallel design solutions. Likewise, Co-3Deator [182] facilitates the sharing and reuse of design-related information.

However, as multiple ideas are materialized in the same physical space (e.g. when two designers add content or annotations around the same object), they need to compete for space, and conflicts arise. Moreover, although some designers can improve their creativity by observing their collaborators' artifacts [247], others may be inhibited by the visibility of their creations [77].

The physical world forces people to perceive a shared workspace in the exact same state. Virtual worlds, in contrast, are not tied to this constraint. When users perceive augmentations through their personal AR equipment (e.g., if they wear an AR headset), they can appear in different virtual spaces. However, both situations have limitations. On one hand, presenting different virtual content to collaborators in the same physical space can isolate them and restrict their collaboration. On the other hand, sharing the same virtual content can limit the number of ideas that can be presented at the same time and may reduce the control that designers feel they have on their own ideas.

In this chapter, I explore a conceptual framework that allows multiple virtual augmentations that may represent alternative design versions to be associated with a single physical object. In this framework, the AR users can control which augmentation(s) they wish to view for a given physical object, independently from their collaborators. They can also create their own augmentations without being constrained by

other augmentations that already exist in a virtual space. The users can also choose to share the designs they have sketched to discuss them with their partners to eventually converge their solutions.

I present a use case scenario that demonstrates the application of this framework. In this scenario, two fashion designers explore ideas by 3D sketching around a physical mannequin. I discuss the concepts of this framework in more detail. I also present a system that implements the key functionality of this scenario.

3.1 SCENARIO

Before I present our conceptual framework in detail, I will start with a scenario that illustrates its main concepts. In this scenario, two fashion designers explore the design of a new female jacket around a physical sewing mannequin. I present it as a creative process expressed through convergent and divergent processes [190].

3.1.1 *Equipment*

Designers are in the same room, in which the sewing mannequin is placed. They are equipped with a Microsoft HoloLens 2 AR headset [266]. They also hold a mouse device with their non-dominant hand that is directly connected to their AR headset. The AR headsets of the two collaborators are connected to the same Wifi network.

3.1.2 *Context*

We¹ assume that our scenario takes place in the workshop of a fashion house. Gabrielle and Karl are two fashion designers who work together on the design of a new jacket. The dimensions of this jacket are constrained by a physical sewing mannequin, located in the workshop. The two fashion designers meet in this workshop for a collaborative design session.

3.1.3 *Simultaneous co-creation*

Karl and Gabrielle both wear an AR headset. Their views are calibrated, such that they can view a common virtual scene at the exact same location. At the beginning of the session, Gabrielle's virtual space (the one she perceives through her headset) is similar to Karl's. They position themselves around the physical mannequin and begin the design process (see Figure 4.9).

¹ This work is the result of a collaboration between my thesis supervisors and myself.



Figure 3.1: **Co-creation with Augmented Reality.** Gabrielle and Karl, two fashion designers, use AR headsets to design a virtual 3D model of a jacket around a physical sewing mannequin.

3.1.3.1 *Creation tools*

Karl and Gabrielle can use tools to interact with their virtual space. These tools are activated by the mouse buttons and then controlled by their dominant hand. A small sphere is displayed at the tip of their index finger to indicate the status and position of the tool. The available tools are as follows:

DRAWING : When this tool is activated, a mesh is generated at the tip of the index finger of the user's dominant hand to form a virtual ribbon (see Figure 3.2.a). The direction of the ribbon is given by the tangent to the path of the index finger. The length of the ribbon is defined by the speed of the finger. This choice is motivated by the work of Dudley et al. [52], who observed that bare-hand drawing appears to be a fast creation tool, making it optimal for quick sketching of ideas.

ERASER : This tool displays a virtual sphere of about ten centimeters in diameter around the tip of the index finger of the user's dominant hand (see Figure 3.2.b). When activated, all ribbon pieces inside this sphere are deleted.

SELECTION : This tool displays a virtual sphere of the same dimensions as the eraser tool. When it is activated, the pieces of ribbon

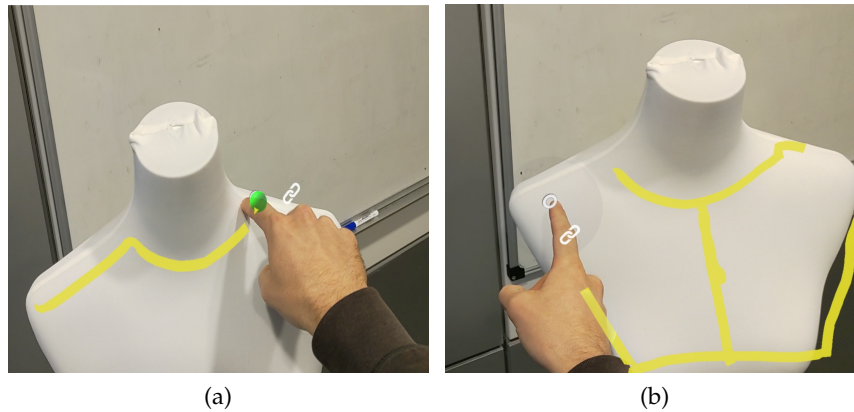


Figure 3.2: **3D sketching tools** for the realization of a virtual 3D sketch on a physical sewing mannequin. (a) Tool for drawing virtual ribbons and (b) tool for erasing virtual ribbons.

that collide with the inside of the sphere are selected and highlighted with an outline effect.

Karl and Gabrielle can open or hide a main menu in the form of a virtual window by pressing the left mouse button. This menu allows them to select a design tool using dedicated buttons. This menu also contains sliders to change the height of the generated ribbons, their hue, saturation, and brightness.

3.1.3.2 Collaborative 3D sketching

Karl draws a virtual 3D line around the mannequin's neck to create a rough collar. As he is not satisfied, he erases the line, and draws a new sketch. Gabrielle perceives the whole set of lines drawn and erased by Karl in her field of vision. While Karl draws his second outline, she approves it verbally and expresses her wish to reproduce a similar curve at the bottom of the cloth. Thus, she draws a 3D curve around the mannequin's waist to illustrate her idea.

To register their progress, Gabrielle creates a version of the 3D drawing through a **Version Object**. To do so, she grabs with her dominant hand the 3D space in which the sketched curves appear (Figure 3.3.a). This gesture allows her to make a Version Object appear, which she then holds in her hand (Figure 3.3.b). Gabrielle moves the object next to the mannequin and then releases it to fix its position in space (Figure 3.3.c). She decides to store all Version Objects that refer to the same design idea around the same location. This allows her to define a **collection** of Version Objects, so that she can later find them easily. Gabrielle now sees the Version Object that corresponds to the current version of the 3D design in her virtual space. She decides not to share it, so Karl cannot see it his own virtual space.

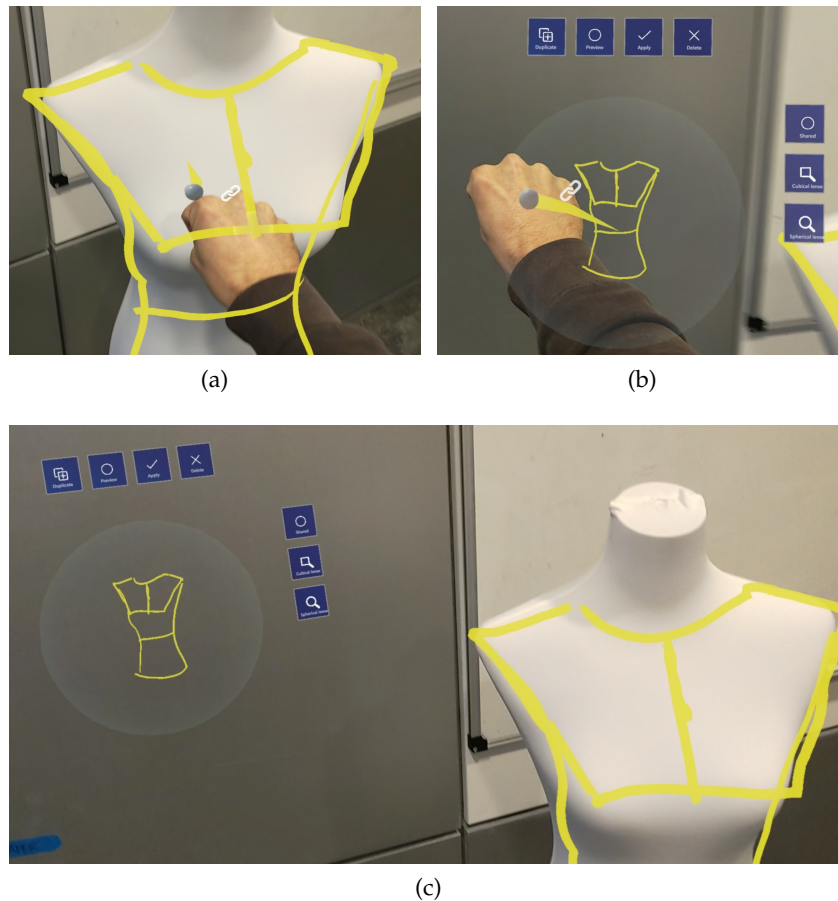


Figure 3.3: **Creation of a Version Object** to save a version of a 3D drawing. (a) Extraction of a Version Object by entering the 3D drawing, (b) moving the Virtual Object in the workspace and (c) storing the Virtual Object in the workspace.

3.1.4 Divergence

As a first step, Gabrielle and Karl generate as many ideas as possible. They produce several 3D sketches of the jacket they want to design to create a rich design space.

3.1.4.1 Desynchronization of virtual spaces

When Karl finishes his new draft of the collar, he notices Gabrielle's tracing of a virtual line around the waist of the mannequin. However, when he approves her idea, he is not satisfied with its rendering. Just as if he would have started a new sketch in his sketchbook, Karl decides to create a new version of the 3D drawing to develop a new idea in parallel with Gabrielle's original idea. In order not to disturb her, he decides to create a copy of the current 3D drawing and to **desynchronize** their virtual spaces. To this end, he creates a Version Object over the current drawing. Several menu options

around the Version Object are now visible (Figures 3.3.b and 3.3.c). He selects the "Apply" item. From now on, he only views the virtual features added to the drawing before the creation of the Version Object. Karl can see Gabrielle working on her version of the drawing, but the strokes she draws gradually disappear. He also notices that an icon above Gabrielle's head indicates that their virtual spaces are not synchronized. The icon above his index (Figure 3.2) has also changed. Karl decides to modify the 3D drawing to further explore his idea, knowing that his changes will not impact the version that is currently viewed by Gabrielle.

3.1.4.2 *Synchronization of collaborators' virtual spaces*

The virtual space of Gabrielle and Karl are now desynchronized, but Karl wishes to take a look at the 3D drawing on which Gabrielle is working. He opens his main menu and makes a "peek" while holding down the corresponding button. By doing so, his virtual space is temporarily put in the state of Gabrielle's space, with a transparency effect on the lines of the 3D drawing. Karl quickly identifies the elements that he wants to discuss with Gabrielle and then releases the button, which returns his virtual space to its previous state.

To fully synchronize his virtual space with Gabrielle's virtual space, Karl clicks the "Synchronize" option in his main menu. Now, he can see the 3D drawing in the same version that Gabrielle sees it. As Gabrielle continues to make modifications to the drawing, Karl sees these changes in real time. Karl discusses the new version of the drawing with Gabrielle and then returns to his original version before the synchronization. He clicks the "Desynchronize" button, which replaces the "Synchronize" button in his main menu. This action desynchronizes their virtual spaces, and Gabrielle can keep working on her version while Karl continues to work on his original version.

3.1.4.3 *Visibility of virtual spaces*

Karl and Gabrielle have desynchronized their virtual spaces. Gabrielle now wants to work on a new idea without Karl seeing it. She decides to click on a button to switch to a **private** mode, making it impossible for Karl to synchronize his virtual space with hers. As a result, the "Synchronize" and "Peek" buttons become inactive.

Gabrielle saves the different versions of her 3D drawing as Version Objects, which are **non-shared**. To allow Karl to see some of the versions she has created, Gabrielle sets the corresponding Version Objects to **shared** by pressing the "Share" button that appears next to the objects. Karl can now see and interact with these objects in his virtual space, in the same positions as Gabrielle.

When Gabrielle is happy with her version, she clicks on a button to switch to the "public" mode and lets Karl know that she has been

working on an idea and would like his opinion. Karl sees the "Synchronize" button become active again in his menu and can click on it to synchronize his virtual space with Gabrielle's virtual space.

3.1.5 *Convergence*

Karl and Gabrielle have completed their ideation phase and are ready to identify the main idea that will guide the realization of the jacket. To do this, they need to compare the different versions of the jacket that they have created and find the most relevant one.

To facilitate this process, they begin by sharing the Version Objects of the jacket that they want to consider. This allows both Karl and Gabrielle to access the same set of objects and compare them side by side.

3.1.5.1 *Browsing alternatives*

Karl is having trouble locating a specific version among the many Version Objects that he has created. However, he does know that some of the features in his current working version are also present in the version he is looking for. To narrow down the search, he decides to use the **filtering by content** function.

Karl begins by selecting certain strokes from his working version. This brings up the existing and shared Version Objects that contain those same strokes. However, he finds that the number of Version Objects returned by the search is too large to easily find the version he is looking for.

To further refine his search, Karl decides to select the strokes from his working version that are not present in the version he is searching for. This operation filters out the Version Objects that share the selected traits and allows him to quickly locate the version he needs.

3.1.5.2 *Visualizing alternatives*

Gabrielle wants to compare her current 3D drawing with another version. To do so, she presses the "Preview" button of the corresponding Version Object to generate a "**preview**" of the other version. This causes the drawing of the other version to be superimposed on her current drawing. Any features of Gabrielle's design that are not present in the other version are highlighted using a set of colors and transparency. Similarly, any features of the other version's drawing that are not present in Gabrielle's current version are also highlighted using colors and transparency.

3.1.5.3 *Comparing and merging different version*

While using the preview of another version, Karl notices a part of a 3D drawing that could be added to his current version. He decides to use the Version Object from this version to generate a spherical **3D portal** (see Figure 3.4). A transparent sphere appears next to the Version Object, and the augmentation as it exists in the other version is visible inside this sphere. The augmentation as Karl perceived it before the portal appeared is visible outside the sphere.

Karl moves this 3D portal and resizes it so that he can see half of the full augmentation in the other version and the other half in his own version. He then selects the features inside the 3D portal that interest him and imports them into his current version. These lines are now present both in the Version Object and in Karl's working version.

3.1.6 *Summary of the scenario*

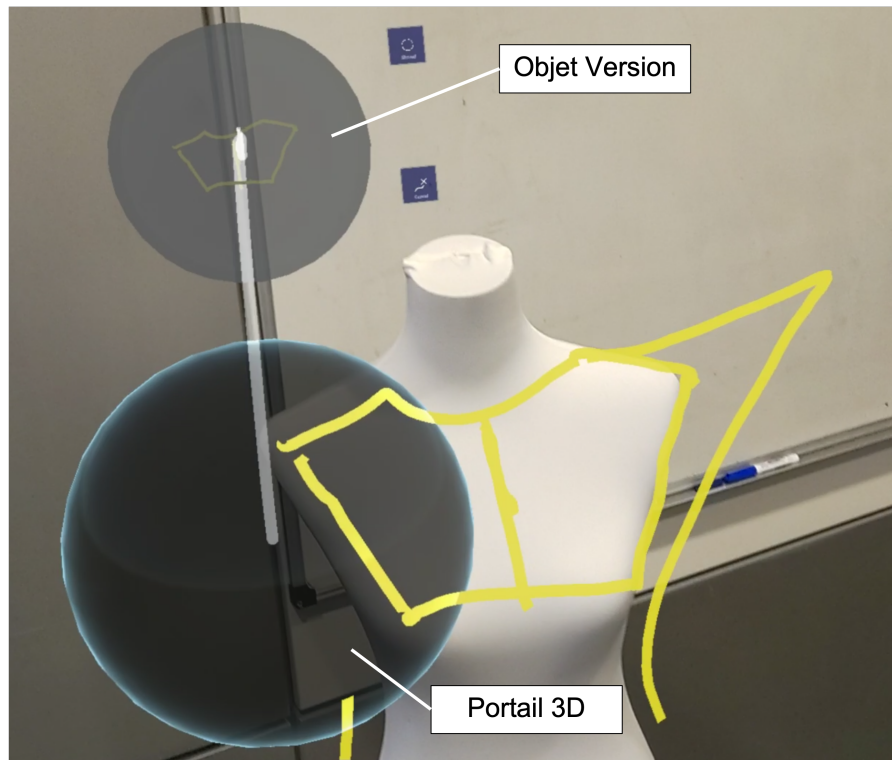
This scenario is an example of collaborative design on a virtual model associated with a physical object. It is based on the use of 3D drawing in Augmented Reality to create sketches or handwritten annotations in 3D. In this scenario, two designers, Gabrielle and Karl, want to save their creation artifacts at specific moments and create, explore, and compare alternatives to these artifacts. They can also use dedicated tools to control the visibility of their creations. The purpose of this scenario is to demonstrate the use of synchronization and desynchronization of collaborators' virtual spaces in Augmented Reality, as well as Version Objects. These concepts are detailed in the following section.

3.2 CONCEPTUAL FRAMEWORK

In this section, I present a conceptual framework that generalizes the scenario discussed in Section 3.1. The framework is designed to address the needs of two collaborators working together in an Augmented Reality (AR) context to design augmentations for a physical object. Each collaborator uses their own AR device to view their own virtual space.

In this framework, a *virtual space* refers to all the virtual content that a collaborator can see. Conversely, the "working space" is defined as the combination of the virtual space and the physical space in which the collaborators are located.

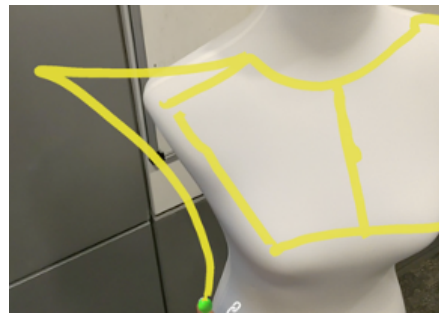
In this framework, the *virtual space* refers to all the virtual content that a collaborator can see. Conversely, the *working space* is defined as the combination of the virtual space and the physical space in which the collaborators are located.



(a)



(b)



(c)

Figure 3.4: **Spherical 3D portal generated from a Version Object.** (a) 3D portal to the virtual space of the Version Object, (b) 3D sketch visible in the virtual space of the Version Object and (c) 3D sketch from the virtual space hidden by the 3D portal

This framework proposes the use of *Version Objects* to version the collaborators' virtual spaces. This allows collaborators to easily switch among different different versions of virtual content, work on the same version, or work on different versions simultaneously by desynchronizing their virtual spaces.

We present a series of tools for collaboration in this context that allow collaborators to create, organize, and compare multiple virtual artifacts simultaneously.

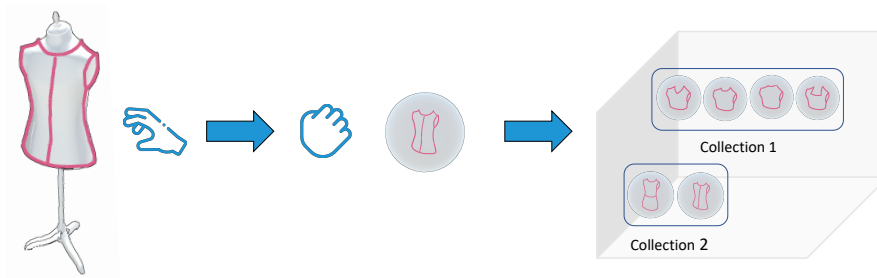


Figure 3.5: **Version Objects creation.** On the left, a version of a 3D sketch overlaying a physical sewing mannequin. In the middle, the corresponding Version Object. On the right, a collection of Version Objects.

3.2.1 Reifying versions

The framework captures the different states of a virtual space at user-defined moments in the form of versions. Version Objects are first-class interactive objects that reify these versions, such that users can easily manipulate their past designs and organize them in space. A Version Object is represented by a virtual object of spherical shape of about twenty centimeters in diameter. It contains a preview of the 3D model that is present in the virtual space, allowing users to identify it among other Version Objects. Its spherical shape enables viewers to observe it independently from their viewpoint, and its size makes it easy to manipulable with a single hand. Users can grab Version Objects and move them in space.

The representation of Version Objects is similar to that of “Containers” in Spacetime [257]. However, the two concepts are different. Containers are projections of a subset of the elements of a virtual space and act as their proxies. They also facilitate navigation in time or in space. In contrast, Version Objects represent the state of a virtual space at a given time, so they facilitate instead navigation among the different versions of the virtual content.

A Version Object is generated based on the virtual information that is currently observable by a user. Its creation can be initiated by an input gesture. A user can then move a Version Object and position it in the workspace. Several Version Objects can be grouped together around the same space to define collections (see Figure 3.5).

3.2.1.1 Creating a version

Like a photograph captured by a snapshot camera, a Version Object depicts the state of a user’s virtual environment at the moment of its creation and is augmented with additional information. By creating Version Objects, users can capture the different stages of their design process or save a 3D model before they edit it to create an alternative. The creation of a Version Object is analogous to the creation of a

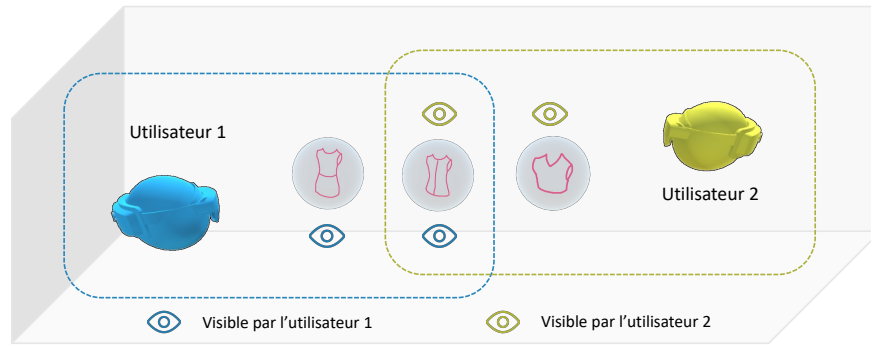


Figure 3.6: **Version Object visibility.** A shared Version Object is visible by all users. A non-shared Version Object is only visible by its creator.

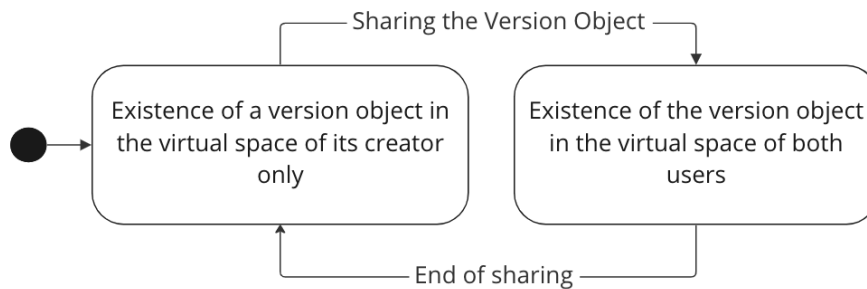


Figure 3.7: State diagram of virtual spaces.

branch in traditional versioning systems [240]. Any new changes to an existing version will be applied to this new branch.

3.2.1.2 *Sharing a version*

When a Version Object is created, only its creator can see it and use it (see Figure 3.6). A user can decide to share a Version Object with other users at any time (see Figure 3.7). Then, the Version Object becomes public, so it is visible and can be manipulated by all users.

This sharing mechanism helps minimize the number of objects that are visible to users. Users can only view the objects they have personally created or have been shared with them by their design partners. Additionally, it enables users to produce a version and privately explore new design paths without exposing themselves to their collaborators' potential judgment.

3.2.2 *Browsing versions*

3.2.2.1 *Switching among versions*

The Version Objects allow users to define and switch among different alternatives for their virtual environment. Each Version Object represents a different version of the virtual content, and users can place

their virtual space in the state corresponding to any of these versions. This enables them to navigate among different virtual environments and explore the design options that have been created.

With the Version Objects, users can easily return to a previous version of the virtual content that they are currently viewing, switch to a different alternative, or create a new alternative if necessary. This functionality provides users with flexibility and control over their virtual environment and allows for efficient collaboration among multiple users.

3.2.2.2 *Visualizing more than one version*

Viewing multiple 3D models simultaneously can be a powerful tool for inspiring and comparing alternative designs. However, when these models overlap, the visualization can become compromised, especially when the 3D models occupy the same physical space. To avoid this issue, the Version Objects in this framework provide solutions by allowing users to generate *previews* and *3D portals*.

Previews. A preview is generated when a user uses a Version Object to inspect the virtual content it represents. This preview adds the elements of the virtual space reified by the Version Object to the user's virtual space (see Figure 3.8.a). The 3D elements that are not present in the version associated with the Version Object used for the preview are highlighted by color and transparency. The same is done for the 3D elements of the Version Object not present in the current version, but with different colors and transparency levels.

This approach enables users to visually distinguish among the common and distinct elements of the different versions. However, occlusion management is not addressed in this conceptual framework. To address this issue, techniques for distorting the space [56] or degrading the 3D models could be used.

3D portals. A 3D portal that is modifiable by the user in terms of its size, location, and geometry can be produced using a Version Object. This portal is based on the use of Magic Lenses [244], whose geometry can be defined by simple volumes, such as cubes and spheres, or more complex with the use of deformable surfaces [151].

A 3D portal creates a "porosity" between the user's current virtual environment and the virtual environment depicted by the Version Object. As a result, it becomes feasible to view 3D elements that are positioned in distinct virtual spaces concurrently while staying in the same physical environment. By moving a 3D portal to an area of the virtual environment where a virtual element from one virtual space coincides with a virtual element from another virtual space, it becomes possible to view two superimposed 3D models at the same time (see Figure 3.8.b). A user can then compare these models by adjusting

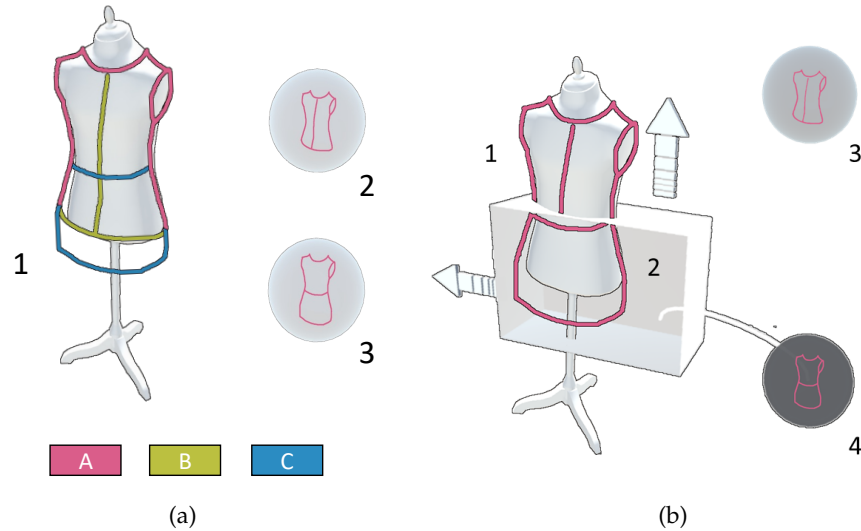


Figure 3.8: Preview (a) and 3D portal shaped as a rectangular parallelepiped (b) to simultaneously visualize two version of a 3D sketching on a physical sewing mannequin. In (a), the 3D sketching in 1 is a preview of the version contained in the Version Objects in 2 and 3. Pink (A) corresponds to the virtual elements shared by versions 2 and 3, yellow (B) to the virtual elements of version 2 not present in version 3 and blue (C) to the virtual elements of version 3 absent from version 2. In (b), the 3D sketching in 1, outside the portal, corresponds to the Version Object in 3. The 3D sketching in 2, inside the portal, corresponds to the Version Object in 4.

the size of the portal, similar to a comparison slider between two overlapping before and after photos.

A preview or a 3D portal can only be generated from a single Object Version and are only visible to the user at the time of its creation. These features are based on the current version of the design and the version contained in this Version Object (see Figure 3.9). Our framework does not allow to generate multiple previews or 3D portals at the same time due to the additional complexity that this functionality would require [244].

3.2.2.3 Searching for versions

To browse the different existing augmentations, a user can select 3D elements present in a virtual space and highlight the objects whose associated virtual spaces do or do not contain these 3D elements. This feature allows users to find versions based on the elements they contain or identify versions whose content differs from the current version.

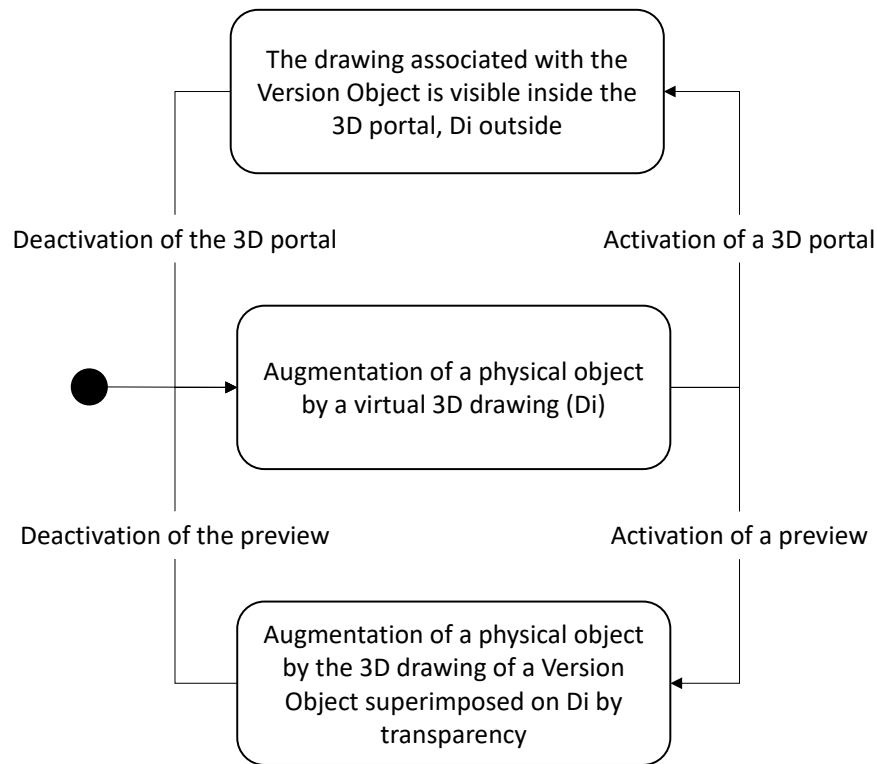


Figure 3.9: State diagram of virtual spaces.

3.2.2.4 Sharing content among versions

Users can import elements from a version into their current version. To do this, they can create a 3D portal using the Version Object containing the version whose elements they want to import and select elements inside this portal to import them. The imported elements are copied into the current version. This operation is analogous to the "merge" operation of traditional version control systems [240].

3.2.3 Collaborating through Version Objects

In a co-located collaboration environment, it is possible to move among different versions by using the previously created Version Objects. However, switching from one version to another is not always a shared decision among collaborators. In order to deal with such situations, our framework allows users to **desynchronize** or **synchronize** their virtual spaces to interact with the same or different virtual spaces. The users have control over the version they own independently of their collaborator. As in TeamStorm [92], a synchronized workspace allows the collaborators to define a group space to work together, while a desynchronized workspace provides a private space for independent work.

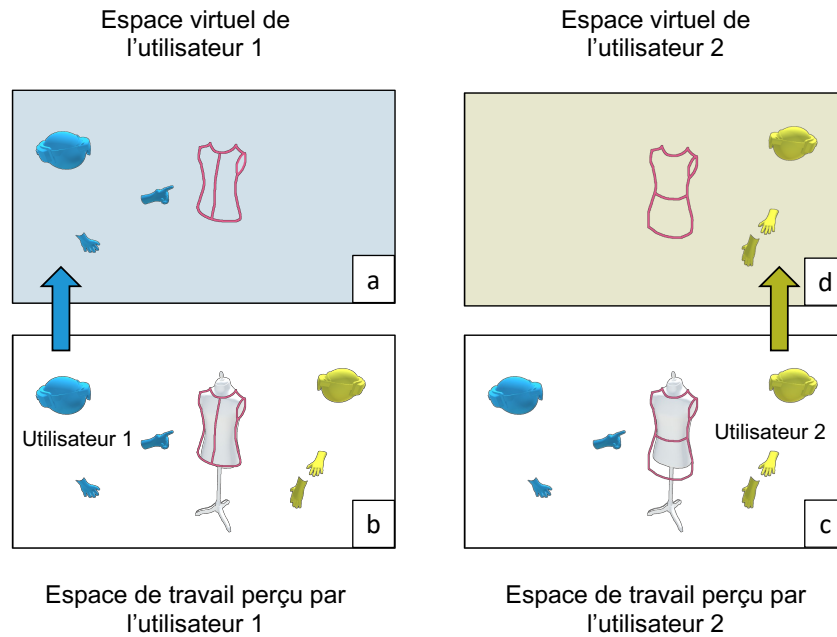


Figure 3.10: **Desynchronized virtual spaces.** Collaborators' virtual spaces are in different states (a and d). They see the same physical mannequin but a different virtual 3D design (b and c).

The desynchronization or synchronization mechanisms do not affect the visibility of the Version Objects. Users perceive the Version Objects they created or are shared with them, irrespective of whether their virtual space is synchronized or not with the one of their collaborators.

3.2.3.1 *Desynchronizing virtual spaces*

Collaborators can desynchronize their virtual space (see Figure 3.10). This operation occurs when a user interacts with a Version Object to switch to its reified version. The virtual space of the user who is willing to switch versions changes accordingly, while the collaborator's virtual space remains the same. This mechanism can, for example, be used to explore different design ideas or to benefit from a private space. When the virtual spaces are desynchronized, users' virtual spaces are no longer synchronized. They move from a metaphor of a shared virtual space, where each user sees and manipulates the same content, to a metaphor of independent virtual spaces, where each user is in a different virtual space.

Then, users do not longer perceive the same augmentation of a physical object. Consequently, further modifications of the virtual content will not be shared. A user can then modify the virtual content associated with a physical object without being constrained by the superposition of other alternative versions.

When virtual spaces are desynchronized, changes made by a user remain temporarily visible in the collaborator's virtual space for a few

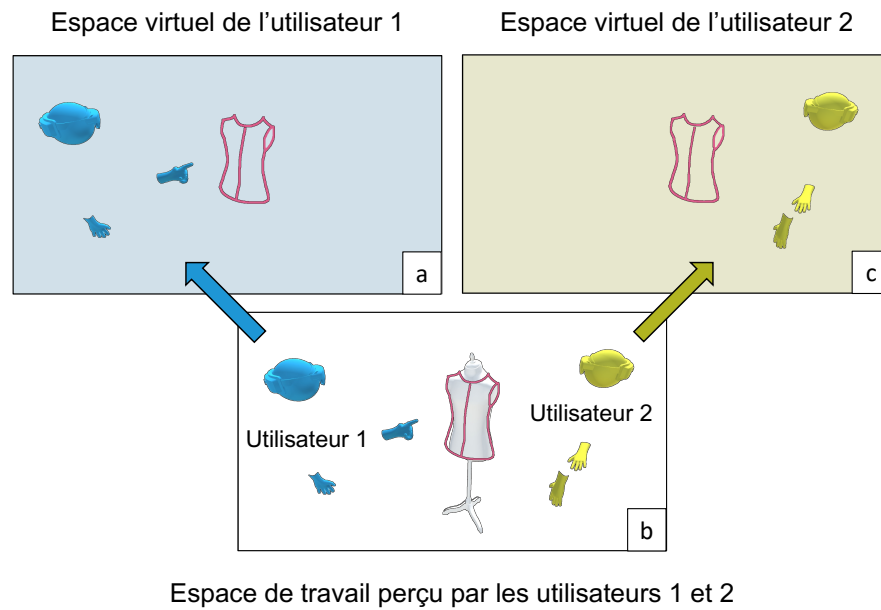


Figure 3.11: **Synchronized virtual spaces.** Collaborators' virtual spaces are in the same state (a and c). The collaborators have access to the same physical mannequin and the same 3D sketches (b).

seconds. This feature provides a temporary overview of the changes made by a collaborator and indicates that the virtual spaces are currently out of sync. However, this temporary persistence can compromise the perception of a collaborator's actions, the design's confidentiality, and the workspace's occupation. Further research is needed to determine the optimal duration of this temporary persistence, taking into account the trade-offs among awareness of changes, privacy, and workspace efficiency.

3.2.3.2 Synchronizing virtual spaces

Users can synchronize back their virtual spaces and want to simultaneously work on the same content or share design ideas (see Figure 3.11). Synchronization happens when a user requests it. The state of this user's virtual space is first saved in a new Version Object before switching to the state of the collaborator's virtual space. This Version Object enables the user to return to the virtual state before the synchronization, if this is necessary.

When users' virtual space becomes synchronized, both users can view and interact with the same virtual content. Changes made by one user become visible to all. In this way, users move from the metaphor of separate virtual spaces to the metaphor of a single shared virtual space.

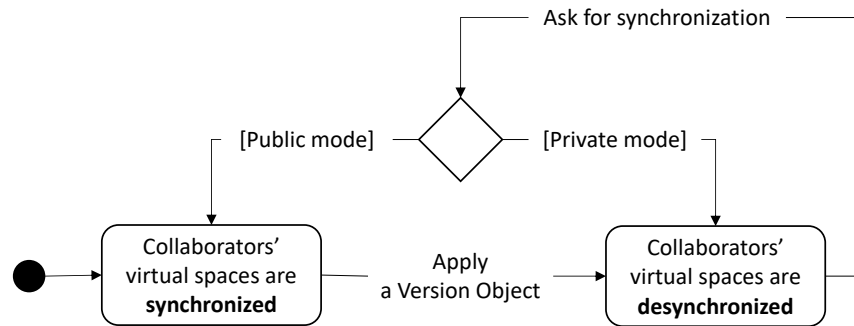


Figure 3.12: State diagram that describes the synchronization and desynchronization of collaborators' virtual space.

3.2.3.3 *Visibility control*

Users can decide not to authorize the synchronization of their virtual spaces to get protected from the interference of their collaborators [77]. As shown in Figure 3.12, a user can choose between a *private* and a *public* mode, where the private mode deactivates the synchronization option. In this state, the collaborators cannot synchronize their virtual spaces with each other. Thus, users wishing to hide their creations from their collaborator, e.g., because they have a work in progress or want to explore an idea without being interrupted, may decide to switch to the private mode.

In public mode, which is activated by default, the synchronization option is active, so the collaborators are free to synchronize their virtual spaces. When both users are in private mode, neither of them can synchronize their virtual spaces with the other. If only one of the users is in public mode, their collaborator can synchronize with them, but they cannot synchronize with their collaborator. These synchronization options are only available when users have their virtual spaces desynchronized.

3.2.3.4 *Viewing the virtual space of collaborators*

When users synchronize their virtual spaces, they get access to the same virtual content. They can then easily understand what their collaborator sees and manipulates.

Users can activate the "**peek**" operation to temporarily switch to their collaborator's virtual space if this is in public mode. The virtual content is then highlighted with transparency to indicate that the virtual space is now identical to the one perceived by the collaborator. This feature is intended to encourage inspiration from the designs of others [247]. It also enables users to preview the space of their collaborators before they eventually decide to synchronize their virtual space.

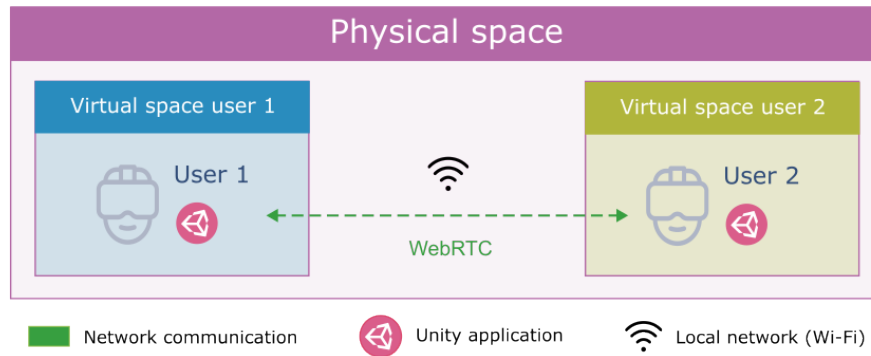


Figure 3.13: System's architecture.

3.3 IMPLEMENTATION

To demonstrate the key interaction concepts and mechanisms of our framework, we have developed a collaborative system that supports the scenario in Section 3.1.

3.3.1 Interaction with the virtual content

The system was developed with Unity 2020.3 [274] and the Mixed Reality Toolkit (MRTK) 2.7 [267]. Each user is equipped with a HoloLens 2 Augmented Reality headset on which a Unity application is run (Figure 3.13²).

3.3.1.1 3D sketch

The 3D drawing is made of ribbons, represented by quadrilateral meshes. The width of the ribbon is given by the user using the main menu. The length of the ribbon is calculated from the speed of the tip of the index finger of the user's dominant hand. The position and direction of the meshes are given by the position and orientation of the tip of the index finger at the time of their creation. These data are obtained via the MRTK and are filtered using the « *1€ filter* » algorithm [27].

3.3.1.2 Version Objects

Version Objects are manipulated with « *Object manipulators* » from the MRTK. They can be moved with a close interaction, by hand, or with a far interaction, by using a virtual ray. Their creation is done by detecting the input gestures (« *grab* ») inside the volume that encloses the 3D drawing (« *bounding volume* »).

² The figure includes icons made by Freepik and Good Ware from www.flaticon.com.

3.3.1.3 *3D portal*

The 3D portal uses the MRTK slicing primitives. Our system automatically updates the list of elements rendered or not through these primitives. The ribbons of the version corresponding to the Version Object used to create it are rendered inside the primitive only. The ribbons of the working version are rendered outside the primitive only.

3.3.2 *Network communication between collaborators*

Our system uses a peer-to-peer network architecture. It connects two Hololens 2 Augmented Reality headsets through a protocol based on WebRTC [275]. Peers are connected via the internet. User-generated ribbons are synchronized in real time between all users. Only the ribbons in the virtual space's current version are visible. This allows for fast rendering during synchronization and desynchronization of virtual spaces and a distributed use of network bandwidth over time.

Only shared objects are synchronized between users. If an object is not shared, its instance exists only in the application of its creator. This minimizes communication between users over the network.

3.4 DISCUSSION

This conceptual framework allows to generate and use several versions of virtual content for the same physical object. I have explored an application case of this framework with a scenario and detailed a possible implementation. In this section, I discuss the scope of this framework and its applications.

3.4.1 *Scope of the framework*

This framework depends on collaborative AR devices where users can perceive a distinct virtual space. Its implementation is thus limited to Augmented Reality systems where each collaborator is equipped with his or her own device for viewing the virtual space. The systems that use projection technologies [49, 105] cannot consequently be used. On the other hand, I consider a collaboration with only two collaborators. However, an extension of this framework to more than two collaborators is conceivable and might require future work.

3.4.2 *Understanding the framework*

The concepts of desynchronization/synchronization of virtual spaces in AR are difficult to represent by analogy with the physical world. This can be an obstacle for users who are not used to AR technologies.

As the complexity of design assistance software can slow down the productivity of its users [140], this framework may also negatively impact some users' creative abilities. Moreover, it is necessary to conduct a user study to understand how the interaction workflows supported by framework can influence collaborative behavior. Will collaborators occupy physical space in a different fashion depending on whether their virtual space is synchronized or desynchronized? How will collaborators make use of Version Objects during the design process? Will they achieve to parallelize design tasks or will they perform them sequentially?

Furthermore, this framework has similarities with text versioning systems, especially for branch creation and merging operations [240]. However, these operations are done through direct interaction with the versioned content. A user can create versions by "grabbing" a 3D drawing. He can also navigate among versions by selecting pieces of 3D drawing. On the other hand, the representation of versions by 3D objects proposes an organization in space that is not addressed by traditional interfaces.

3.4.3 *Implications for the design process*

The use of a computer causes interruptions in the design process [141]. This is due to the sequences of operations required to perform a task, often performed with a keyboard and mouse (pointing, menu navigation, ...). By manipulating 3D objects to perform operations of saving or comparing digital content, this framework aims to reduce the incidence of these interruptions and thus make the design process more fluid. The development and the study of systems based on this framework would permit the evaluation of its incidence and compare it to the design support tools currently used.

3.4.4 *Extending the framework to Virtual Reality*

Virtual Reality makes possible interactions that are not possible in a physical environment. It allows, for example, to use different scale factors between collaborators, duplicate content, or navigate in space by teleporting [257]. However, it immerses its users in complete virtuality, which prevents them from directly perceiving their physical environment. Even if they can be complex and expensive to implement, reconstruction techniques can circumvent this limitation by including all or part of a physical environment in a virtual space [34, 219]. Although the perception of the physical environment is simpler in Augmented Reality, the use of Virtual Reality in this framework could extend its functionality.

3.4.5 *Extending the versioning mechanisms to mixed-reality objects*

This framework aims to facilitate the generation of alternatives to virtual content associated with a physical object. In this approach, however, I consider this physical object as immutable. Bousseau et al. [18] suggest that designers often explore ideas and discuss solutions through direct manipulation or transformation of physical materials, often in combination with drawing. In some design scenarios (e.g., neighborhood planning by urban planners), this may translate into the need for collaborators to manipulate or modify physical objects (e.g., models of urban furniture). The capture of these manipulations and transformations in Augmented Reality can be done through tracking techniques [260], reconstruction techniques [34, 219], or using intelligent materials [253]. An evolution of this approach would be to use these techniques to extend versioning to mixed objects. The versions could then represent the state of the physical space in addition to the state of the virtual space.

AR technologies radically change the way 3D design teams work together. AR users can move away from the screen of their computer to interact directly with the objects of a virtual scene and naturally navigate in their physical space. AR also strengthens collaboration by adding virtual aids [170] while preserving traditional communication channels, such as voice, gaze and gestures. Previous work has investigated the use of AR for a diverse range of collaborative tasks, from interior design for couples [208] and science teaching [223] to industrial manufacturing [251]. Unfortunately, real-time collaboration is a challenge when users work remotely and, consequently, they do not share the same physical environment and do not all have access to AR equipment. Such situations have become commonplace during the still ongoing COVID19 pandemic [272]. Many design and research teams have found themselves to work remotely, relying on video-communication software to collaborate together [258]. Some experts predict that such situations are not temporary – they will largely persist after the pandemic [15]. HCI research thus needs to better understand how different remote workspace configurations support collaboration in these new contexts.

While screen sharing has been a valuable tool of collaboration for remote desktop users, sharing the workspace of a collaborator wearing an AR headset requires a new set of tools that considers both the physical and the virtual space of the AR user. In this direction, several AR technologies, such as the Microsoft HoloLens enable AR users to video-stream their view. Yet, such views are not interactive and do not offer independent camera control to remote viewers. According to Tait and Billingham [227], increased view independence results in stronger collaboration performance. However, view independence requires that the physical environment of the AR user is reconstructed in real time, such that it can be smoothly integrated into the 3D virtual scene. Unfortunately, existing solutions for reconstructing independent AR views have serious limitations. For example, techniques based on multiple depth sensors [1, 8] require heavyweight instrumentation, consume large volumes of bandwidth, while the quality of their reconstructed models is still limited and largely unrealistic [110]. Other 3D reconstruction techniques [74, 163] pose significant constraints on the view possibilities of remote users.

The alternative approach that we investigated in this work is to offer remote users multiple view representations, where each provides a different aspect of the workspace of the local AR worker. We focused,

in particular, on tasks that require access both to a virtual model and to its physical context, or to physical objects that interact with the virtual model. In this case, remote collaborators must make decisions about which representation to use to effectively complete the task.

We studied three complementary representations: (i) a first-person view as provided by the AR headset, (ii) an augmented third-person view as captured by a fixed camera with a depth sensor, and (iii) a fully virtual representation. The first two representations show the physical-world scene but do not support view independence. The last representation, in contrast, supports full view independence but does not capture the physical-world scene. However, by providing tools for switching among these representations, we expected that remote users would develop strategies that leverage their complementary roles. We framed our research questions as follows:

RQ₁: How do remote users perceive the trade-offs of the three representations when providing instructions to an AR worker?

Several past studies [67, 75, 223] have studied the trade-offs of first-person and third-person views, but as we discuss in this paper, their results are somehow contradictory and non-conclusive. Others [227] have studied fully independent views, but only ones that rely on the 3D reconstruction of the physical scene.

RQ₂: If we offer remote users the possibility to switch among representations, how will they make use of them? To explore answers to this question, we integrate the three representations into ARGus¹ (see Figure 4.5), a remote collaboration system. A key contribution of ARGus is on how its user interface merges representations through a collection of interactive tools for pre-viewing, between- and within-view navigation, camera control, 3D pointing, and annotation.

In the following sections, I will report on the results of two user studies, one for each question. The first study examines strengths and weaknesses of the three representations, focusing on the collaboration experience of remote users when communicating spatial instructions. The second user study investigates how 12 remote participants use ARGus to guide a local AR user to complete an AR furniture arrangement task. Our results provide a fresh perspective on the trade-offs of each representation. They also help us characterize participants' view-switching strategies, evaluate the perceived effectiveness and utility of ARGus, and understand whether and how it assists remote communication.

¹ <https://argus-collab.github.io/>

4.1 DESIGN PROBLEM

We are interested in asymmetric collaboration setups that involve a *local user* with an AR headset (e.g., a Microsoft HoloLens) and *remote collaborators* who participate from distance through a desktop application. In contrast to approaches that require users at both ends to wear an AR or a VR headset [8, 239], such setups are relatively lightweight and easy to employ, as they only require the local user to have access to AR equipment. These setups thus offer high flexibility to the *remote collaborators*, allowing them to work in many different situations, such as while traveling or in a crowded open office where physical space is limited.

Video has become the most common medium of remote collaboration and has taken a dominant role during the ongoing COVID19 pandemic [258]. Our goal is not to replace video communication but to enhance it with new visual and interaction modalities that leverage the benefits of AR systems. A major challenge is how to deal with the asymmetry in the views of remote collaborators, in particular how to enable them to easily navigate in the 3D environment of the AR user, inspect the virtual content, and provide directions that require spatial orientation and awareness.

As we already discussed, we also dismiss solutions that require the reconstruction of the physical workspace [8, 111, 227, 235], either because they cannot keep track of dynamic changes in the environment of the local user, or because they provide a largely unrealistic representation of the scene and the local user, break the collaborators' experience due to the "*uncanny valley of XR telepresence*" [110], and amplify network outage problems [3]. We restrict our design space to lightweight configurations that use a single external depth camera in addition to the camera of the AR user's headset. This external camera could be replaced by a webcam or a smartphone since more and more devices are now equipped with a depth sensor. We may even rely on standard webcams or smartphones in the near future, as a single monocular camera can be sufficient to provide depth data [153].

Focusing on the views of the *remote collaborators*, we investigate three design dimensions:

WORKSPACE REPRESENTATION. It refers to the representation used by the system to help collaborators perceive each other and their shared workspace. This representation may consist of a virtual 3D scene, video, or alternatively a combination of these two. Ideally, it should provide spatial information about both virtual and physical objects in the workspace but also information about the actual AR user, such as her or his body position and gestures.

SCENE VIEWPOINT. It determines which virtual and physical objects are visible at a given moment or during the whole collaborative

task and from which perspective. Previous literature often makes a distinction between a *first-person* and a *third-person* perspective (e.g., see Komiyama et al. [129]). The former refers to the perspective of the AR user. It can be captured by a head-mounted and communicated to remote collaborators. The latter refers to an out-of-body perspective as captured by external cameras.

VIEW INDEPENDENCE. A key problem is how to enable remote users to independently navigate in the 3D space of the AR user to obtain a convenient view, e.g., a view that helps them inspect details of the virtual model or avoids occlusions, and point to a position in space, e.g., to indicate a physical or virtual object to the local user.

Additional dimensions, such as display configuration and means of communication, can emerge from this design space. We chose dimensions that focus on the collaboration process itself rather than ones that deal with how collaboration is made possible, since many tools have already been presented for this purpose [75, 223, 235, 254]. To simplify our user studies, we also decided to focus on one-to-one collaboration. We defer the study of the more general case where multiple remote collaborators participate to our future work.

Since we do not expect the remote workspace reconstruction problems to be solved any time soon, we limit our scope to techniques of augmented video-mediated communication, as those require more lightweight setups, consume less bandwidth, and do not suffer from 3D reconstruction problems. Furthermore, as we study tasks that involve both virtual and physical objects, we are also interested in how streamed video can be coupled with fully virtual representations that afford free navigation.

4.2 TRADE-OFFS BETWEEN VIEWPOINT, WORKSPACE REPRESENTATION AND NAVIGATION CONTROL

We conducted a user study to investigate our first research question (RQ₁). The study examines trade-offs of different workspace representations and scene viewpoints. In particular, it observes how users provide remote instructions under three configurations:

HEADSET VIEW is an augmented video from a first-person viewpoint.

We capture the video directly from the AR headset to simulate the situation where the remote user sees the scene “through the eyes” of the local user (see Figure 4.1-a). The video feed integrates the virtual 3D content into the physical scene of the AR user. The key strength of this configuration is that collaborators share a common frame of reference. So they do not need to mentally rotate the 3D space [207] to communicate.

EXTERNAL VIEW is an augmented video from an external (third-person) viewpoint. We use a depth camera (Microsoft Kinect V2) to provide an overview of the full workspace of the local user. A key question is how to optimally position the camera. In previous studies [67, 223], in which the local worker remains seated, the external camera is positioned at the back left (or right) side of the worker. This way, the two collaborators view the scene from a similar perspective. Unfortunately, in such configurations, the face, hands, and other key parts of the worker's body may not be visible. Furthermore, if the worker freely moves around the model of interest, his or her body may occlude parts of the workspace. For these reasons, we excluded this alternative. For optimal visibility, the camera is positioned in front of the AR user (at 2m height and approximately 2.5m away) and oriented 30° downwards. We also ensure that the board on which the local user places objects is centered in the recorded image. The video feed of this camera is augmented with the virtual 3D content visible in the AR user's workspace (see Figure 4.1-b). Compared to first-person views, external views have been shown to increase communication efficiency [67] and improve performance and satisfaction [223]. Other authors observe that users strongly prefer them for "*placing objects recommended by themselves*" [22].

VIRTUAL VIEW is a fully virtual representation with a free viewpoint. Remote collaborators see a virtual representation of the 3D scene. A simplified avatar shows the head and hands of the AR user (see Figure 4.1-c). Remote users can freely navigate in the 3D scene and choose their preferred viewpoint. This approach follows the naive metaphor of birds that can choose the most convenient position to observe the AR user. Previous results [227] suggest that this additional freedom in the choice of views can improve both the performance and the confidence of remote collaborators.

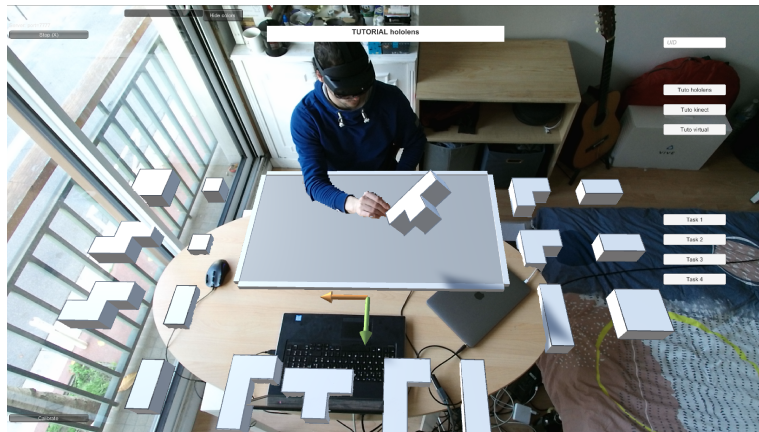
The study took place during the COVID19 pandemic. To eliminate risks of contamination, the experimenter (first author) acted as the local user wearing the AR headset for all sessions of the study. Participants acted as remote collaborators and completed the study tasks from their home or office environment. The experimental protocols of our studies were approved by a local ethical committee.

4.2.1 Participants

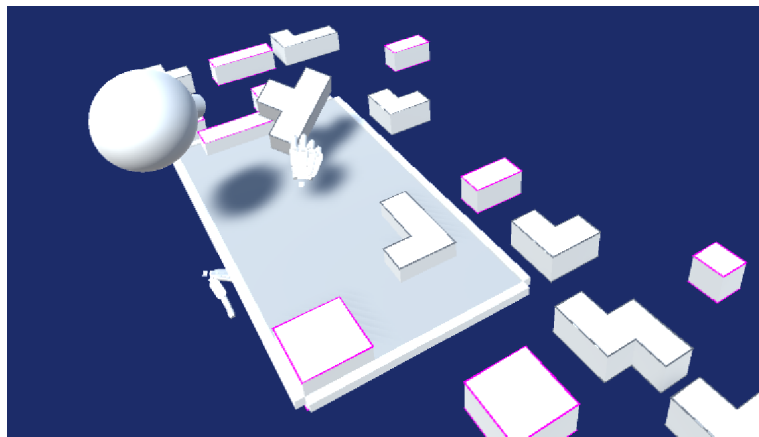
24 volunteers (11 women and 13 men) participated. They were 21 to 41 years old (Median = 26.5 years). All were frequent or occasional users of at least one video-communication tool, such as Skype or Zoom. Seven participants frequently or occasionally used an AR or



(a)



(b)



(c)

Figure 4.1: Remote-view configurations tested by our first study: HEADSET VIEW (left), EXTERNAL VIEW (middle) and VIRTUAL VIEW (right). A remote participant gives oral instructions to the AR user on how to position 3D shapes on a virtual support.

a VR headset. 11 participants were frequent or occasional users of 3D games, game engines, or 3D modeling environments. Participants

were recruited by word of mouth and responses to a recruitment email sent to our lab's mailing lists. No compensation was given.

4.2.2 *Apparatus*

The experimenter set up the workspace in his home environment and interacted with the scene through a Microsoft HoloLens 2. For the calibration, the experimenter defined the HoloLens origin by manually positioning a 3D object on an AprilTag [250] marker. The Kinect camera was automatically calibrated by detecting this marker using the ViSP library [156]. Communication between the participants and the experimenter was established through commercial video-communication software (Skype or Discord). The HEADSET VIEW and EXTERNAL VIEW were presented to participants through screen sharing. For the HEADSET VIEW, we used the Microsoft HoloLens 2 video-sharing application [268] to stream live video from the headset. For the EXTERNAL VIEW, our implementation considered potential occlusions between virtual and physical objects as "seen" by the Kinect camera. For each pixel, a shader chose to display either the streamed video or the virtual object by respecting their depth information from the camera. For the VIRTUAL VIEW, participants downloaded and executed a client application, which rendered an interactive 3D scene synchronized with the HoloLens application via a remote server. This architecture is implemented with Unity 2019.4 and used the Unet library [273] for network communication. Participants could pan, zoom, and rotate the 3D scene using their mouse and keyboard. Finally, we used a website to guide participants in the course of the experiment (see Figure 4.2-a). This website provided information and instructions regarding the configurations and the task and linked to our online questionnaires.

4.2.3 *Task*

Participants were asked to place 3D pieces of nine different shapes on a virtual board by giving oral instructions to the experimenter who acted as a surrogate (see Figure 4.1). The experimenter used close and distant manipulation tools provided by the Microsoft HoloLens 2: its direct manipulation gestures, its hand-ray tool and air-tapping for selection.

The solution to the task was a 2D top-view pattern that described how to position pieces in any order. The pattern was randomly generated to contain eight pieces out of 18 pieces available in the workspace. It was presented to participants as an image on the website and was unknown to the experimenter (see Figure 4.2-b). Its default orientation shown to participants reflected the experimenter's perspective. The pattern was thus inverted with respect to the EXTERNAL VIEW. To help participants adapt the orientation of the 2D pattern as they would do

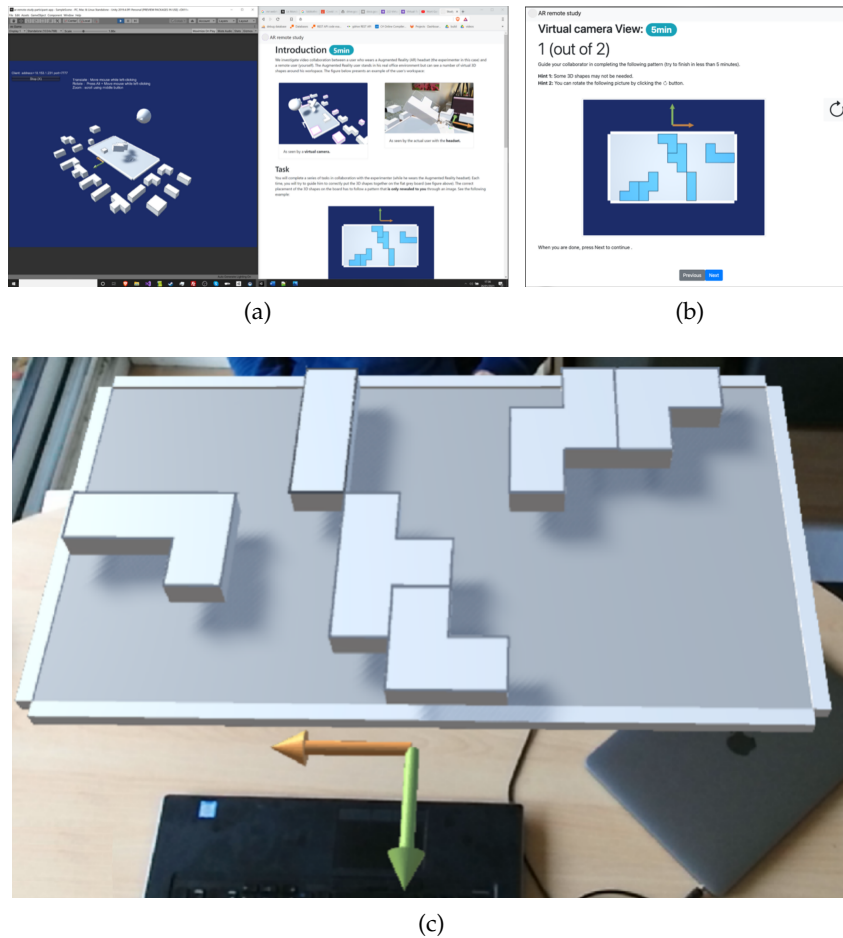


Figure 4.2: (a) Remote participant interface used for our first study: tested view configuration on the left (VIRTUAL VIEW in this example) and website used to give instructions on the right. (b) Close-up of the website showing the target pattern: the UI widget on the right allows participants to rotate the pattern image. (c) Zoom-in on the AR user workspace showing the virtual board with the finalized task: colored axes help participants make the correspondence between the pattern on the image and the virtual board shown on the view.

with a piece of paper, we included UI widgets for rotating the pattern. We also added colored axes both in the views and pattern images to make correspondence clear. The virtual board was composed of a 9×5 grid of squares with side length 10 cm (see Figure 4.2-c). When a 3D piece was placed on the board, it was snapped to the grid. Pieces had a maximum length of 30 cm (i.e., three grid squares).

As Kuhlen and Brennan [131] discuss, using a confederate in studies that involve conversations between humans is a common research method, but its practice “*might be hazardous*” to collected data. In particular, if confederates have an active, uncontrolled participation in the dialog and are aware of the hypotheses of the study, they can bias

the results. To reduce the risk of bias, we established a minimalistic communication protocol for the experimenter. The experimenter followed the participant's instructions and only verbally intervened: (i) to ask the participant to repeat an instruction if the instruction was not understood; (ii) to request confirmation for a planned action; and (iii) to request confirmation for a completed action. The experimenter could also answer questions concerning the user interface or the task, but we tried to respond to such questions as much as possible during training. In contrast, the task required participants to take the initiative as speakers, as Kuhlen and Brennan [131] also recommend.

4.2.4 *Design*

To keep sessions short, we simplified the experimental design by dividing the user study into two independent parts. Focusing on the viewpoint (first-person vs. third-person) of AR video, PART I compared the HEADSET VIEW with the EXTERNAL VIEW. Focusing on the workspace representation (virtual vs. AR) and the type of navigation control (remote user vs. local user control), PART II compared the HEADSET VIEW with the VIRTUAL VIEW. We divided our participants into two groups of 12 participants, one for each part, trying to balance gender. We followed a within-participant design, where all 12 participants tested both configurations. Half of them were first exposed to the HEADSET VIEW, and the other half starts with the second condition. For each configuration, participants completed two main tasks, preceded by a training task with a simplified pattern with three only pieces.

4.2.5 *Procedure*

After signing a consent form, participants completed an online demographic questionnaire. Participants went through a short tutorial that explained the two communication configurations. They were then introduced to the training and two main tasks of each configuration. Participants evaluated the configurations and the task through a set of questions divided into multiple short questionnaires. Each participant answered seven questionnaires in total: one after each task (2 tasks \times 2 configurations), one after each configuration (2 configurations), and one after the full session. The full procedure lasted approximately 50-70 minutes.

4.2.6 *Data Collection and Measures*

We collected: (i) participants' answers to the online questionnaires, (ii) recordings of the participants' voice during the tasks, and (iii) logs of low-level software events (view positions, trajectories, and time

stamps). As we discussed above, the presence and collaboration role of the experimenter adds bias in the way tasks are completed. As a result, task performance measures, such as task-completion time and errors are not reliable, and we do not consider them here. We focus instead on how participants perceived difficulty for different components of the task. We also report on the participants' preferences and their feedback about trade-offs of the compared conditions. Finally, we examine the strategies that participants followed to complete the tasks. Consider that our analyses are exploratory and should be interpreted as such.

4.2.7 Results

We present our main results. Anonymized data from this study and the R code of our analyses are available as supplementary material at <https://osf.io/g7xas/>.

Perceived task difficulty. Participants rated the difficulty for each sub-task through 5-point Likert items (1 = very difficult, 5 = very easy). We miss the answers of one participant for these questions in PART I. The analysis of ordinal data with metric models is generally problematic [146]. We therefore use state-of-the-art *cumulative probit* regression models [25, 146] that enable us to map ordinal scales to a latent (i.e., not observable) continuous variable and then express estimates of differences between conditions as standardized effect sizes. For an extensive justification of this method and a comprehensive tutorial, we refer the interested reader to Bürkner and Vuorre [25]. The method is based on a Bayesian statistics [119] framework, but we emphasize that we do not use informative priors here.

Figure 4.3 presents the results of our analysis, where we compare the perceived difficulty of our configurations through estimates of mean standardized differences expressed as 95% *credible intervals*². Those are differences over a continuous (rather than ordinal) physiological variable of difficulty and are expressed in standard deviation (SD) units. In contrast to common non-parametric significance tests that rely on rank transformations, the approach enables us to estimate the magnitude of the observed effects by means of probabilistic interval estimates and effect sizes and thus better evaluate the statistical evidence about these effects.

The results indicate that participants perceived that the EXTERNAL VIEW was easier than the HEADSET VIEW for searching pieces in their collaborator's environment. In contrast, the EXTERNAL VIEW was more difficult for describing how to translate or rotate a piece and how to

² A credible interval is the Bayesian analog of a confidence interval. Unlike a 95% confidence interval, which is often misinterpreted, a 95% credible interval expresses a range in which the parameter of interest lies with 95% probability [119].

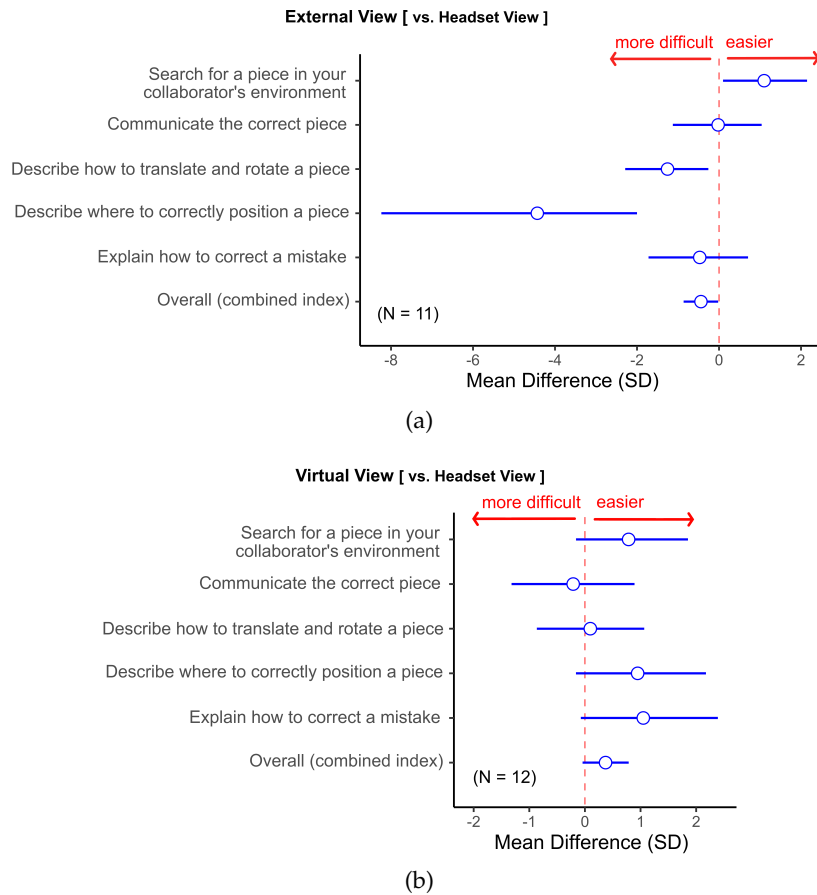


Figure 4.3: Comparing the perceived difficulty of different subtasks among configurations. For our analysis, we use Bayesian ordinal (cumulative probit) models [25], which map the original ordinal scale of Likert items to a *latent* continuous variable. The bars in the graph represent 95% *credible intervals* of mean differences over this continuous variable and can be treated as estimates of standardized effect sizes. Note that the unit of these differences is the standard deviation (SD) of the distribution of the latent variable.

correctly position a piece. This latter effect is especially pronounced. When exposed to the EXTERNAL VIEW, several participants struggled to correctly map their image of the pattern to the workspace of the AR user. Because of the position of the external camera, the participants had to mentally perform a rotation transformation to give the correct instructions. We further discuss this problem below. For the other subtasks (communicate the correct piece and explain how to correct a mistake), we do not observe any clear difference between the two configurations.

Differences between the VIRTUAL VIEW and the HEADSET VIEW are more uncertain. There is a trend that the VIRTUAL VIEW was perceived as easier for searching pieces in their collaborator's environment, for describing how to correctly position a piece, and for explaining how

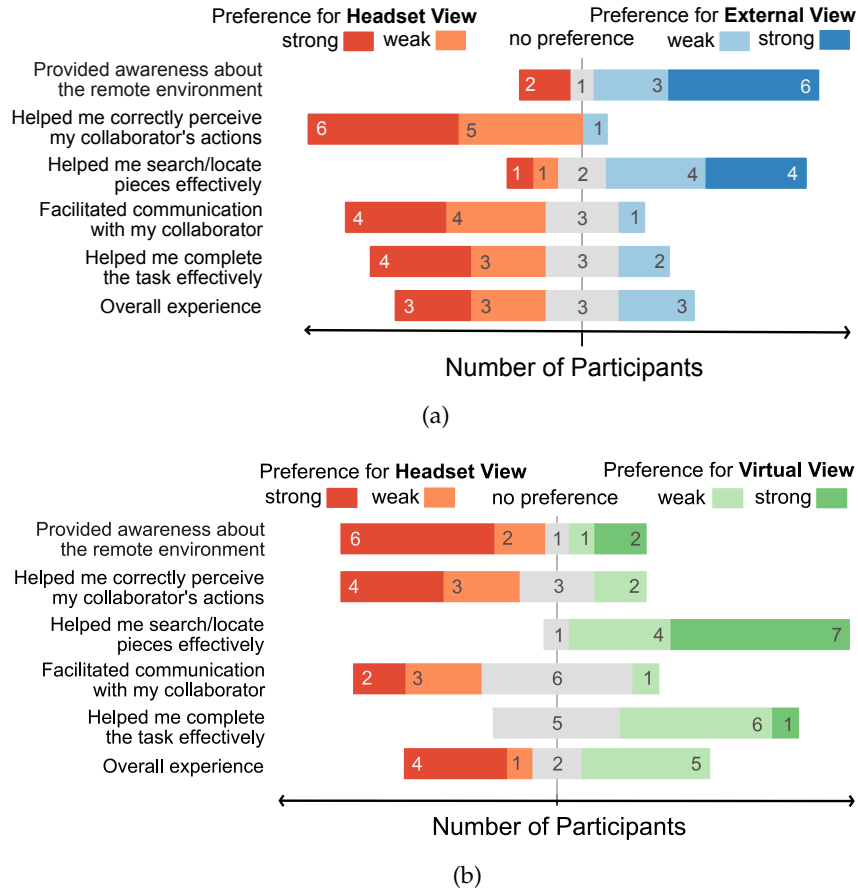


Figure 4.4: Distribution of participants’ preferences: HEADSET VIEW vs. EXTERNAL VIEW (left) and HEADSET VIEW vs. VIRTUAL VIEW (right).

to correct a mistake. However, the low size of the sample does not let us draw clear conclusions.

Preferences. We also asked participants to compare the configurations that they tested on six different aspects of the collaboration task. Figure 4.4 summarizes our results. We observe that participants see different benefits in each configuration. They appreciated the ability of the EXTERNAL VIEW to provide awareness about the remote environment and help them search and locate pieces effectively. However, most participants expressed an overall preference for the HEADSET VIEW, as it helped them perceive their collaborator’s actions, facilitated communication, and helped them complete the task more effectively. The VIRTUAL VIEW, in turn, was especially appreciated for helping participants search and locate pieces effectively but also complete the task more effectively than the HEADSET VIEW. Overall preferences between the VIRTUAL VIEW and the HEADSET VIEW were equally split.

Trade-offs. Open-ended questions in the questionnaires asked participants to elaborate on the strengths and weakness of each configuration. All 12 participants of PART I reported that providing a global

view of the workspace was the main strength of the EXTERNAL VIEW. *"The strongest aspect was being able to see the overview of the scene and the entire puzzle we are building as a whole"* (P1). *"The fixed camera implies that all the items always stay in view of the distance person, easier if the collaborator cooperates less"* (P11). As a comparison, in the HEADSET VIEW *"the environment is reduced, and it takes more time to find your way around and locate all the items"* (P3). *"I do not have an autonomy of my vision angle, I only see what he sees"* (P5).

However, most participants evaluated this very same property of the HEADSET VIEW as its strongest aspect: *"giving directions is much easier because I can just tell the partner to what I am doing!"* (P5). According to P12, *"you see through the eyes of [your partner], so you could exactly guide his gestures like a puppet."* In contrast, eight participants explicitly mentioned the inversion of left and right as a major problem of the EXTERNAL VIEW: *"You are located on the opposite side so everything is going to be the reverse to explain."* (P12). Even though we allowed users to rotate the reference image with the solution pattern (see Figure 4.2-b), only half of them used this function, and even this strategy did not seem to solve the problem for them. As additional limitations of the EXTERNAL VIEW, participants complained about distance distortions (P10), difficulties in correctly perceiving depth (P1), a sense of *"distantiation"* (P12), and a weaker sense of participation (P3).

The responses of the participants of PART II focused on the same qualities and drawbacks of the HEADSET VIEW but raised additional concerns that the camera can be *"shaky"* (P19) and can *"induce motion sickness"* (P13). Concerning the VIRTUAL VIEW, participants especially appreciated its navigation capabilities: *"The user may navigate independently of the operator, make it possible to change point of view, or see things out of the operator's sight"* (P13); *"you are totally autonomous on the vision of the environment"* (P21). However, participants also identified several weaknesses: *"I do not really know where my collaborator is looking at"* (P15); *"lack of information about the real environment of the other user"* (P18); *"less points of reference than the previous configuration"* (P22); *"users need to be used to 3D applications in order to place [their] view correctly"* (P20).

Communication strategies. All participants frequently referred to their partner's "left" and "right" to communicate orientation. A common approach for indicating specific objects was to verbally describe their shape, e.g., by means of a letter of a similar shape ("Z", short "L", long "L", etc). A small number of participants (four in total) responded that they sometimes or frequently made use of physical objects in the experimenter's space as reference for the two AR views. To provide directions about how to rotate objects, strategies were more diverse. Several participants described the angle (90 or 180 degrees) of the rotation and its direction (clockwise/anticlockwise or left/right), while two participants acknowledged difficulties in finding an efficient strat-

egy. For translations, most participants used the edges and corners of the virtual table for reference, but for higher precision, they also referred to the borders of other pieces on the table. In the VIRTUAL VIEW, participants' dominant approach was to place the virtual camera above the head of the avatar of their partner to obtain a similar viewpoint. According to our logs, four participants moved around the board to discover a better viewpoint but also ended up placing the camera at this position.

User feedback. Two participants proposed to place the camera of the EXTERNAL VIEW slightly behind (P5) or above the head (P12) of the AR user, while P2 proposed to approach the camera closer to the table. P14 and P21, instead, wondered about the possibility to increase the field of view of the HEADSET VIEW, e.g., by adding extra cameras, while three participants (P12, P13, P15) proposed to combine multiple views together. Finally, several participants made suggestions about pointing techniques: a cursor for "*indicating locations*" (P17), a laser to "*target specific parts*" (P18), clicking with the mouse to "*illuminate a piece*" (P22) or to "*ping*" at a certain position as the HEADSET VIEW moves (P16), and "*add a vocabulary to easier describe pieces*" (P9).

4.2.8 Discussion

The task required the AR user to manipulate virtual only objects. This choice was made to ensure that participants could complete the task under all three configurations. Clearly, it overrates the utility of the VIRTUAL VIEW, which lacks support for physical objects. Furthermore, we notice that some participants expressed strong preference for the HEADSET VIEW over the EXTERNAL VIEW. The EXTERNAL VIEW was also rated as more difficult for certain subtasks. This finding is somehow at odds with results of past studies [67, 223], suggesting that the specificities of the task and the camera viewpoint may have an important influence to the success of a representation. In particular, in the external views that those two studies compared, the camera was conveniently located to the back left of the worker. As Shepard and Metzler [207] have shown, the time needed to perform a *mental rotation* in 3D space linearly increases with the angular offset of a viewer's viewpoint. This mental-rotation model predicts longer reaction times for our 180° camera configuration and implies a greater mental effort. An 180° offset also requires collaborators to reverse their wording, e.g., to replace every egocentric "right" with a "left" [204].³

Despite the above shortcomings, the EXTERNAL VIEW presents several benefits over the HEADSET VIEW. First, the view provided global awareness about the remote environment. Second, most participants

³ A mirror configuration would transfer the problem to rotational directions, e.g., a "clockwise" direction should become "anticlockwise." Given their complexity, we suspect that the mental effort of such transformations would be even greater.

felt that it helped them search for and locate pieces with less effort (see Figure 4.3-4.4). The EXTERNAL VIEW is also the only configuration that allows remote users to see the face and physical full body of their collaborators. Although the role of such information was not directly evaluated with our task, it can be essential for supporting empathy [228] among participants and establishing communication awareness [89].

4.3 *ARGUS*: A MULTI-VIEW COLLABORATION SYSTEM

The results of our first study show that each view configuration has unique qualities that are difficult to substitute by the other two. The EXTERNAL VIEW supports global awareness about the physical environment of the local worker and helps the remote user search for objects that are spread around the workspace. The VIRTUAL VIEW supports independent navigation, helping the remote user to provide instructions (e.g., about how to correct mistakes) from a convenient but also stable point of view. Finally, the HEADSET VIEW is especially effective for perceiving the actions of the AR user and communicating egocentric instructions. Our research efforts thus focus on how to combine them and how to give remote desktop collaborators direct control over their use. To this end, we developed ARgus, a multiview collaboration system for 3D modeling. ARgus's implementation reflects three design goals:

- dg₁**. Communicate both real and virtual representations but without requiring the 3D reconstruction of the local workspace. We rely instead on video for capturing the physical environment of the local AR user and his or her physical body. As we discussed in previous sections, this approach avoids problems associated with the 3D reconstruction of a physical workspace.
- dg₂**. Support both first-person and third-person views of varying levels of view independence. This goal is consistent with the results of our formative study and recommendations of several older studies [75, 111, 173, 198, 227]. A challenge for ARgus was how to design effective and consistent mechanisms for switching and navigating between and within views.
- dg₃**. Provide tools that minimize communication effort and facilitate coordination. According to Schober [204], speakers try to minimize the mental effort of their addressees and their own by replacing *speaker-centered descriptions* (e.g., at "my left" or "your right") by *neutral descriptions*. ARgus provides aids for neutral descriptions via direct-pointing and spatial-annotation tools.

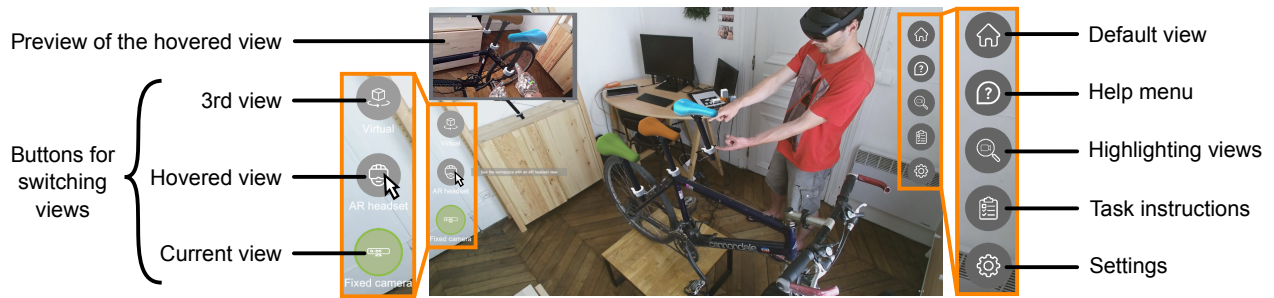


Figure 4.5: Desktop interface of ARGUS used by a remote collaborator for the redesign of a bicycle saddle.

Below, we present the main features of ARGUS. Although the system supports bidirectional communication, we focus in this paper on its design for remote desktop collaborators.

4.3.1 Combining Multiple Views

ARGUS receives the augmented video streams from both an AR headset and an external depth camera located in the AR user’s physical space. Furthermore, it maintains a synchronized version of the virtual 3D scene and can generate virtual views from any workspace location. Remote users can seamlessly switch between virtual and augmented video representations, as well as freely navigate to any viewpoint on the 3D scene.

ARGUS also offers the possibility to display live previews of all three views (HEADSET VIEW, EXTERNAL VIEW, and VIRTUAL VIEW). These previews are video thumbnails of alternative views displayed in a small embedded window on top of the user’s current view. They allow users to take a quick look at a different view, e.g., to inspect details of the physical environment that are not visible in the current view or to decide whether it is worth switching views. This mechanism aims to prevent the short bursts of switching between views observed by Gaver et al. [75] and facilitates coordination when users ask their collaborator to temporarily switch to their viewpoint to approve the veracity of their discovery [173].

4.3.2 Supporting Navigation

We provide several solutions for displaying previews, switching between views, and navigating in the 3D scene.

Main user interface. The main window of ARGUS’ user interface displays three circular buttons for selecting views and getting feedback about the active view (see Figure 4.5). When users hover over a button, a live video preview is displayed on the top-left corner of the window. Clicking on the button activates the view. We use

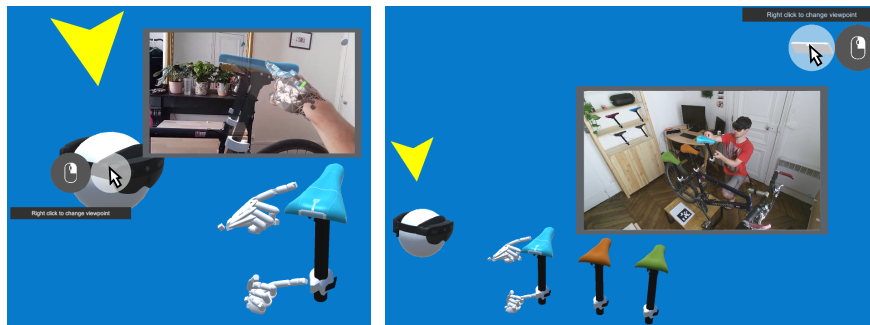


Figure 4.6: The remote user hovers the mouse over the headset of the 3D avatar (left) and Kinect 3D model (right) to display the preview of the HEADSET VIEW and the EXTERNAL VIEW respectively.

a trajectory and field-of-view interpolation based on Cinemachine [265] to animate the virtual camera in the 3D scene. This solution ensures visual consistency among views, helps users understand the location of distant viewpoints, and avoids disorientation. We also use a blur effect to smooth out transitions between augmented video and virtual representations. We let users customize the duration of view transitions.

Interacting with the 3D scene. The 3D scene of *ARGUS*' VIRTUAL VIEW serves as the basis for 3D navigation. It also offers an alternative solution for switching between views through interactive virtual camera representations. In the VIRTUAL VIEW, users can use the mouse to rotate their viewpoint around the center of the 3D scene and translate it (pressing ALT). The same navigation capabilities are available in the two augmented-video representations, the HEADSET VIEW and the EXTERNAL VIEW. However, since remote users do not have direct control of the position of the two physical cameras (i.e., the external and the head-mounted camera), navigation actions within these views immediately cause the view representation to turn to virtual. This design approach ensures that interaction is consistent across all views.

The 3D scene includes virtual representations of the physical cameras themselves. Users can interact with them to preview or activate their corresponding views. For example, Figure 4.6 shows the active VIRTUAL VIEW of a desktop user who remotely collaborates for the redesign of a bicycle saddle. The virtual view does not provide information about the physical scene. Therefore, the remote user hovers the mouse over the headset of the 3D avatar to better understand what her partner sees (Figure 4.6-left). She then hovers over the model of the Kinect camera (Figure 4.6-right) to compare how the three saddle designs look together with the bicycle's physical frame. Users may also decide to click the mouse to switch to this view. Finally, the 3D scene includes guides (arrows and highlighting effects) that help users locate the cameras and orient themselves in the 3D space.

Navigating with spherical views. Using base 3D rotation and translation interactions to closely inspect specific parts of a 3D model can be tedious and time consuming. To facilitate such tasks, we adapt *Navidget* interaction technique [90] and integrate it into *ARGUS*' user interface as a *SPHERICAL VIEW* tool. Activated with a mouse right-click within either the *EXTERNAL VIEW* or the *VIRTUAL VIEW*, the tool visualizes a sphere centered on the selected point. Users can move a virtual camera on the surface of the sphere, and a camera preview is shown (see Figure 4.7-left). The sphere radius can be adjusted with the mouse wheel, causing the virtual camera to zoom in or out. Users can release the mouse to switch to a desired view or press *ESC* to keep the current viewpoint.

Viewpoint recording. Following the approach of Sukan et al. [222], we allow users to record viewpoint locations (pressing a key) when they spot interesting views that they want to later reuse. Viewpoint recordings are represented as virtual cameras. As all other cameras (see above), they have a visual representation in the 3D scene, and users can interact with them to preview or switch to their views.

4.3.3 *Facilitating Communication*

Other tools in *ARGUS* focus on how to facilitate the communication of users (*DG*₃).

AR user representation. The *VIRTUAL VIEW* includes a synchronized representation of the AR user with a simplified avatar composed of a sphere wearing the 3D model of a Microsoft HoloLens 2 and virtual hands (see Figure 4.6). Each hand is represented by 24 joints, connected by canonical shapes, such as cylinders and squares. Both hands and head positions are retrieved from the MRTK libraries [267]. In the *EXTERNAL VIEW*, a vertical arrow on top of the physical head of the AR user communicates an interaction point for previewing and selecting the *HEADSET VIEW*. If view point representation is not visible through the current viewpoint, arrows are pointing on the side of the screen in their direction to easily retrieve them. This technique have also been used, for example in [143] or [235].

Pointing stick. As several participants of our formative study proposed, it is often useful to directly point in the remote scene, e.g., to indicate an object or provide instructions about where to place it. *ARGUS* provides such functionality through a *VIRTUAL STICK* (see Figure 4.7-middle). The stick starts from the viewpoint's origin. Its direction is controlled with the mouse, while its length can be adjusted with the mouse wheel. A small sphere represents its tip, which is red if colliding with a 3D element and grey otherwise. A dotted line indicates its pointing direction, starting from its tip and projected until

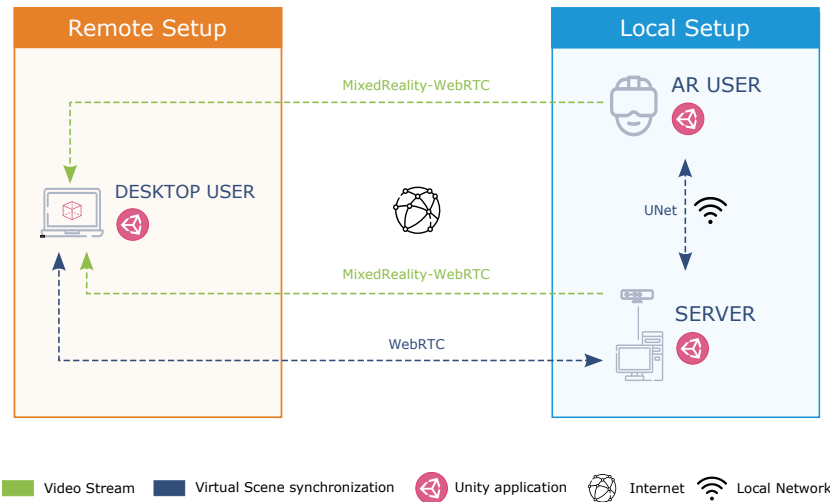


Figure 4.8: System's architecture and implementation.

crease mental workload [229]. Likewise, using shared virtual landmark increase user experience and facilitate spatial referencing in collaboration [165, 204]. In all views of ARGUS, remote users can use the VIRTUAL STICK to add annotations represented as colored spheres. In Figure 4.7-right, for example, the remote user has added a yellow and a blue annotation to suggest target locations for placing furniture. The user interface shows a list of all activate annotations (up to five in our evaluation study), allowing users to quickly review and remove them.

4.3.4 Architecture and Implementation

ARGUS was developed in Unity 2019.4. Its architecture relies on a client-server model connecting a remote desktop user and a local AR headset to a local server (see Figure 4.8⁴). The server keeps a synchronized version of the 3D scene and records the AR user's physical workspace with the external depth camera. It generates the EXTERNAL VIEW by augmenting the camera video feed with the objects of the 3D scene. Occlusions between the virtual objects and the physical objects are managed through the depth map of the external camera: for each pixel, a shader displays either the streamed video or the virtual object according to their depth information.

The AR headset is connected to the server as a client using the Unet library. It maintains a synchronized version of the 3D scene, which is used both to render AR user's augmented view in the headset and to generate the video feed of the HEADSET VIEW. It also transmits the AR user's head and hands positions and orientations. To calibrate the AR headset reference frame and the depth camera reference frame, the virtual space origin is defined (i) manually by the AR user who

⁴ The figure includes icons made by Freepik and Good Ware from www.flaticon.com.

needs to position a 3D object on an AprilTag [250] marker and (ii) automatically by the depth camera, which detects this marker using the ViSP library [156].

The application of the remote user is also connected to the server as a client using WebRTC. We built a custom protocol based on this technology to synchronize 3D object states (position and rotation) and software events (tools, logs, etc.). The application can thus render a synchronized version of the 3D scene to create the VIRTUAL VIEW. In addition, it receives the video feeds from the AR headset and the server based on the Mixed-Reality WebRTC libraries [269] to display the EXTERNAL VIEW and the HEADSET VIEW.

4.4 USER STUDY 2

We conducted a second user study that investigates our second research question (RQ₂). The study examines how remote collaborators use ARGus to provide instructions to a local AR designer.

As for our first study, we opted for an experimental design that avoids contamination risks due to the COVID19 pandemic. The experimenter (first author) acted as the local user wearing the AR headset, while participants acted as remote collaborators and completed the tasks from their home or office. A preregistration [38] of the study is available at <https://osf.io/6dhzn>.

4.4.1 *Participants*

12 volunteers (4 women and 8 men) participated in the study with an age ranging from 24 to 29 years old (Median = 27.5 years). All were frequent or occasional users of at least one video-communication application. Two participants frequently or occasionally used an AR or a VR headset, while five participants had no previous experience with AR/VR technologies. Eight participants were frequent or occasional users of 3D games, game engines, or 3D modeling environments. Before starting the tasks, we verified that all participants had a stable internet connection (we replaced four initial participants who could not continue due to connection problems). We followed the same recruitment process as for our first study.

4.4.2 *Apparatus and Conditions*

As for the first study, the experimenter interacted with a Microsoft HoloLens 2 in a workspace created in his home environment. We evaluated a simplified version of the ARGus (written here as ARGus) to help participants quickly master the key features of the interface. More specifically, we deactivated its support for viewpoint recording since it was not useful in our experimental task. We also used pre-



Figure 4.9: Experimenter wearing a Microsoft HoloLens 2 (a) collaborating with a participant using ARgus on a desktop computer (b).

selected positions for the spherical view, suitable for the 3D model used in this study. To activate the tool, participants had to right-click on a yellow cylinder located at three relevant positions of the model (one for each room of a house model). The cylinder then became the rotation center of the SPHERICAL VIEW. As we observed in our first study, finding a good placement for the external depth camera is not trivial and largely depends on the task. We decided to use the same configuration as for the first study: we positioned the camera at 2m height, 30° downwards to face the experimenter and to capture his moving body and his augmented workspace, minimizing occlusions. We used the HEADSET VIEW as control condition. As in our first study, this condition did not provide any interaction capabilities.

Participants downloaded and executed a single Unity application for both conditions on their personal computer. The user interface had a fixed-size window with a 1920×1080 resolution. A step-by-step tutorial about the system functionality and the tasks was directly embedded in the system. For verbal communication between the participants and the experimenter, we used a commercial application (Skype or Discord).

4.4.3 Task

ARgus' functionalities can support remote mixed-reality participatory design in a range of domains, such as furniture arrangement [22] and urban planning [31, 200]. We decided to focus on a furniture arrangement task because it was used in the past by other related studies [75, 111, 201]. As in our formative study, this task requires participants to search for 3D pieces in the workspace of the experimenter, find a target location for them, and instruct the experimenter to place them correctly. In contrast, we now looked for tasks that would involve both physical and virtual objects in a scene. We considered two alternatives: (i) the AR user manipulates virtual pieces within a larger physical

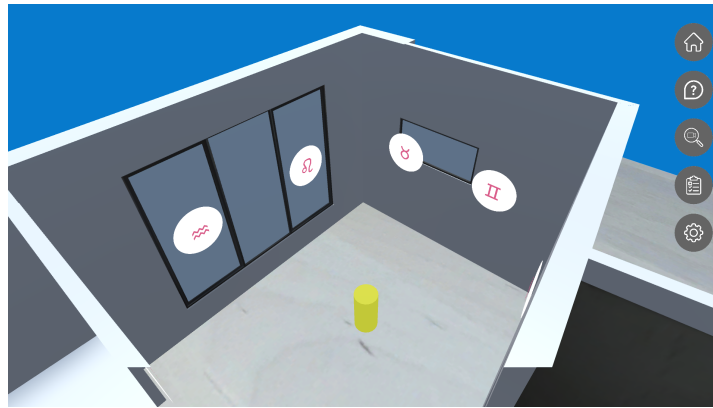


Figure 4.10: Symbols on the walls of the virtual house model to constrain the placements.

frame of reference (e.g., as in Figure 4.5); or (ii) the AR user manipulates physical pieces (miniature furniture) within the virtual model of a house. We opted for the second alternative, as it provides richer opportunities for virtual navigation and better captures the trade-offs of different representations. The task simulates the situation where a remote buyer communicates with a furniture designer (or seller). The furniture designer follows instructions to try miniature models of his or her collection in a virtual model of the buyer's house.

We introduced several constraints to create various arrangement tasks unknown to the experimenter. Zodiac symbols were randomly displayed on pre-defined positions on the virtual house model's walls (see Figure 4.10). We chose these symbols as they are easy to identify but hard to verbally describe. This way, we forced participants to rely on intrinsic landmarks of the model for communicating positions, rather than artifacts that are absent in real-world tasks. Two symbols were randomly assigned to each participant. In each room, these two symbols were located on perpendicular walls and defined a cross-shaped forbidden area: the line in front of each symbol was not available to place furniture.

Participants were asked to arrange furniture for three *thematic spots* randomly chosen among nine. The functional aspect of these spots was described textually. For example, a "living spot" was described as "a place where people can meet and spend some time together". To perform this task, participants could choose miniature furniture among six storage cabinets, four tables and ten chairs (see Figure 4.11).

To complicate the task, we required each miniature chair to be appropriately oriented so that sitting people can see a virtual window without moving their head too much. To be valid, a spot had to include at least two pieces of furniture, meet the placement constraints and represent an harmonious layout (according to the participant's preferences). The symbols, constraints and spot description were communicated to participants at the beginning of each task and made



Figure 4.11: Miniature furniture manipulated by the experimenter according to the participant's instructions.

available at any time in a specific panel of the interface (see Figure 4.5). This information was unknown to the AR local user.

As for Study 1, we tried to reduce the experimenter's influence [131] by constraining his verbal interventions to only ones required for the completion of the task, i.e., asking the participant to repeat instructions, and asking for confirmation of planned or completed actions.

4.4.4 *Design and Procedure*

We followed a within-participant design, where all 12 participants tested both user interface configurations. Half of them were first exposed to the HEADSET VIEW. The other half were first exposed to ARGUS. After signing a consent form, participants completed an online demographic questionnaire. They were then introduced to the two configurations. For ARGUS, participants went through a tutorial presenting each tool step-by-step. For each configuration, participants completed a practice and main task. The practice task required the arrangement of one thematic spot.

At the end, participants completed a questionnaire that evaluated their experience with the two configurations that they tested. The full procedure lasted approximately 70 to 90 minutes.

4.4.5 *Data Collection and Measures*

We collected participants' answers to a pre- and a post-questionnaire. The post-questionnaire evaluated the efficiency of each user interface configuration with a Likert scale of four items with seven levels (1 = *Inefficient*, 7 = *Efficient*). It also assessed the importance of verbal communication for each configuration with a Likert scale of four items with five levels (1 = *Not important*, 5 = *Very important*). The

questionnaire further evaluated the utility of the views and interactive tools of ARGUS configuration and collected participants feedback about their use. We also collected logs of low-level events that describe the use of interactive tools and view transitions during the task. Due to technical problems, logs were not collected for one participant (P5).

Finally, we recorded and manually transcribed participants' voice during the tasks. We then distinguished among phrases that provide remote instructions and other non-instructional content, such as transitional ("ok", "now") and thinking-aloud sentences. Instructions were further classified into three subtask categories: identifying & reaching an object, manipulating an object, and moving in the scene. These categories cover the full set of instructions that we identified and do not overlap. We started with a finer-grained coding scheme. In particular, we initially tried to differentiate among instructions on identifying and reaching objects or locations, and among instructions that concerned different types of manipulation actions. However, these categories were often fused, which made their coding uncertain and unreliable. We thus finally opted for larger categories.

The first and second author decided together on how to segment the transcripts and code the segments by inspecting the data of the first participant. They independently coded the transcripts of three additional participants. They then discussed and finalized the segmentation and coding scheme. As a last step, the first author re-coded all the transcripts, while the second author independently coded the transcripts of the last two participants. We calculated inter-coder reliability at the word level both for distinguishing between instructions and non-instructions (Krippendorff's $\alpha = .98$, 95% CI [.97, .99]) and for the overall classification that also considers the type of instruction (Krippendorff's $\alpha = .97$, 95% CI [.96, .98]). Inter-coder reliability scores are high, so we count and analyze the words in participants' transcripts for all the above categories.

4.4.6 Questions and Hypotheses

We expected that participants might develop diverse strategies to complete the tasks. Our goal was to observe and understand these strategies. We were particularly interested in two questions:

Q₁: Will participants find the three views of ARGUS useful, and how will they make use of them?

Q₂: Will participants find the user interface tools useful, and how will they make use of them?

Furthermore, we wanted the participants to reflect about how they completed tasks with the two configurations and report on their trade-offs. Tait and Billinghamurst [227] found that increased view independence reduces the number of verbal instructions among collaborators.

Likewise, we expected that ARGUS would reduce reliance on verbal communication, because it gives more viewing freedom to remote users and provides opportunities for completing the task more efficiently. More formally, we tested the following three hypotheses:

- H₁: The mean perceived efficiency will be higher for ARGUS.
- H₂: The mean perceived importance of verbal communication will be lower for ARGUS.
- H₃: The mean number of words for communicating instructions will be lower for ARGUS.

Like Tait and Billinghamurst [227], we are interested in the link between view independence and communication performance. However, our studies are distinct from each other. First, since we do not reconstruct the model of the physical scene, we investigate view independence through complementary views with different levels of navigation control. Therefore, we also try to identify the view-control strategies that participants develop to carry out the task. Second, our system includes an external view, which also shows the physical body of the local user. Note that Tait and Billinghamurst [227] recognize the potential benefits of an external view and identify it as a promising configuration for future studies. Third, Tait and Billinghamurst [227] test the positioning of physical objects on a physical table. We study instead a more complex task that requires collaborators to position physical pieces within a larger virtual model. In our case, collaborators need to deal with occlusions in the AR scene, thus both physical and virtual navigation are essential for completing the task. Finally, annotations in their system are virtual replicas of a small collection of physical objects, which are conveniently placed on the surface of a table. Our annotation mechanism is simpler but more generic, as it lets remote participants mark any virtual or physical object and location in the 3D workspace with little manipulation effort.

4.4.7 Results

Anonymized data from this study and the R code of our analyses are available as supplementary material at <https://osf.io/g7xas/>. Here, we summarize our results.

Use of tools and view representations. We first summarize the strategies that participants used to complete the task under the ARGUS condition. For each participant (except for P5), Figure 4.12 visualizes the active views during the task and the use of previews, the pointing stick, and the spherical view. We emphasize that we did not encourage participants to be fast, and the time range that we show does not always reflect active collaboration time. Some participants (e.g.,

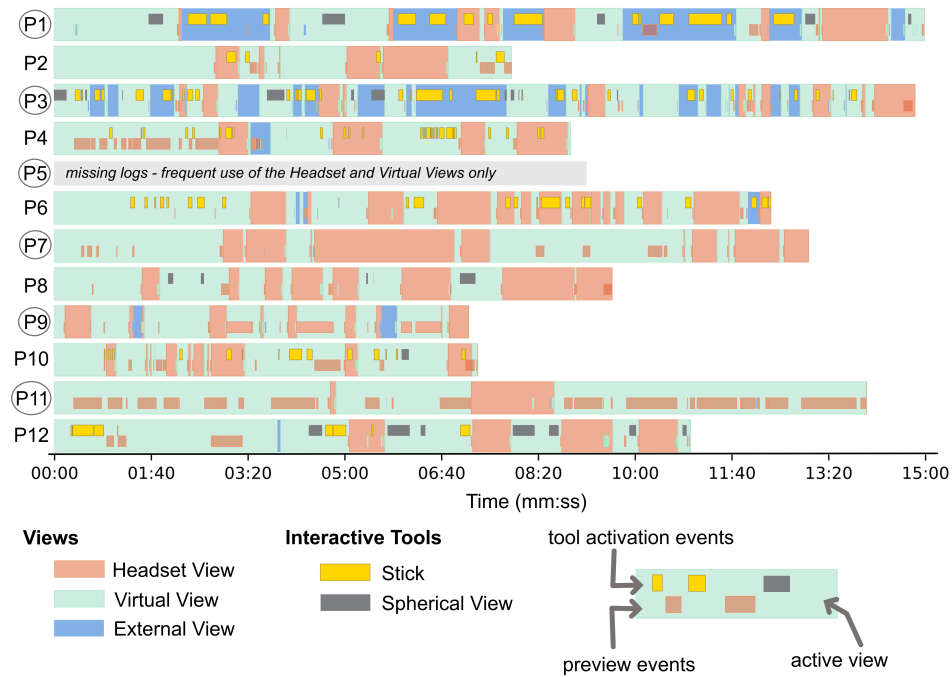


Figure 4.12: Use of the three view representations, the pointing stick, and the spherical view by the participants of the evaluation study for the main task under ARGUS. Circled participants were exposed to ARGUS first.

P1) spent initial time to think about the constraints of the task and further explore the available tools. It is not a surprise that the slowest participants in Figure 4.12 were exposed to ARGUS first (in circle).

Overall, all participants frequently transitioned among views during the task, which demonstrates the utility of our approach. However, we observe that the VIRTUAL VIEW and the HEADSET VIEW dominated the participants' choices. The EXTERNAL VIEW was heavily used by P1 and P3 and sparingly by three other participants. Participants' questionnaire responses are consistent with these patterns. Three only participants found the EXTERNAL VIEW to be useful (P3) or very useful (P1, P9). P2 explained that he did not *"feel the need"* to use it but *"in a bigger environment it could have been useful to guide the partner quickly from one point to another."*

The three view representation were used in two different ways: (i) as main active views or (ii) through the preview window. Figure 4.12 shows that several participants (P2, P4, P7, P9, P10, and P11) extensively used the HEADSET VIEW in preview mode from the VIRTUAL VIEW. According to P4, *"the headset view caused dizziness [...] I stayed in the virtual view and watched the headset view from the window."* P2 agrees that *"having the headset view showing in the corner while navigating and pointing in virtual view was the ideal setup."*

The stick was activated in all three representations either as a pointing or as an annotation tool. For example, P₁ and P₃ regularly used it from the EXTERNAL VIEW to indicate furniture pieces. P₄, P₆, and P₁₂ used in combination with the VIRTUAL VIEW to indicate target positions. Other participants did not feel the need to use it: *"I did not use the stick as the rooms had enough identifiable elements to allow my partner to understand my instructions"* (P₅). Finally, a smaller group of participants made use of the spherical view. According to P₁, it is *"the best to manage the constraints"* but other participants did not agree: *"I was comfortable enough with virtual navigation not to feel the need to resort to the spherical view"* (P₂); *"I tried to use the spherical view but I am not enough comfortable with in comparison with rotate and translate so I abandoned."* (P₆); *"I would have liked a 2D mapping"* (P₅). The spherical mapping that we used is generic but may not be the most appropriate for the specific task. Alternative mappings that better adapt to the geometry of the virtual model might indeed improve the usability of the tool.

Perceived efficiency. We compare the efficiency of the two user interface configurations as perceived by our participants. We use again Bayesian cumulative probit models [25] for our analysis (see Section 4.2). Figure 4.13 summarizes our results. Overall, participants rated ARGUS as more efficient (see Hypothesis H_1). This was especially the case for verifying the constraints in the scene. For this task, free navigation through the virtual view seemed to be crucial. According to P₆, the HEADSET VIEW causes *"seasickness"*, while P₉ commented that its resolution *"was not so effective to perceive accurately the symbols on the walls when having a wide point of view."* In contrast, seven participants rated the HEADSET VIEW as more efficient for helping them to perceive the workspace of their partner despite the fact that the ARGUS configuration provided a richer set of views and options for observing the remote space. The added complexity of this interface can explain this result: *"Having only one solution forces to rely on it and in the case of the headset, forces to establish an efficient communication with the partner, that can be lacking when overwhelmed by all the possibilities of the different views and the difficulty to master them all"* (P₃).

Reliance on verbal instructions. Figure 4.14 compares the mean difference between configurations in participants' perception about the importance of verbal communication. Overall, verbal communication was perceived as less important for ARGUS (see Hypothesis H_2), particularly for describing which pieces of furniture to take and where to place them. P₁₀ explained that verbal communication is more important for the HEADSET VIEW *"because you cannot point with as much precision as with the stick and you cannot see equally well symbols and distances."*

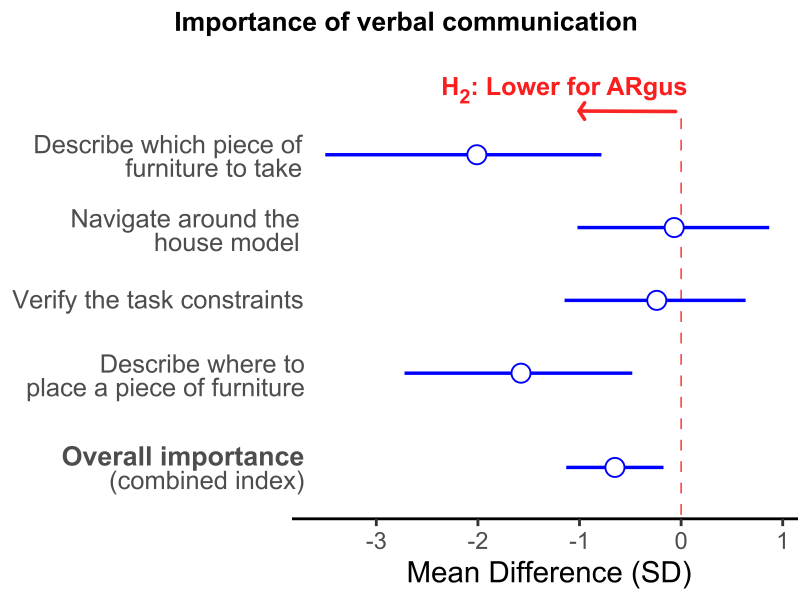


Figure 4.13: Comparing the perceived efficiency of the two user interface configurations ($N = 12$). We use again Bayesian ordinal (cumulative probit) models [25]. The bars in the graph represent 95% credible intervals of mean differences over a latent continuous variable and can be treated as estimates of standardized effect sizes.

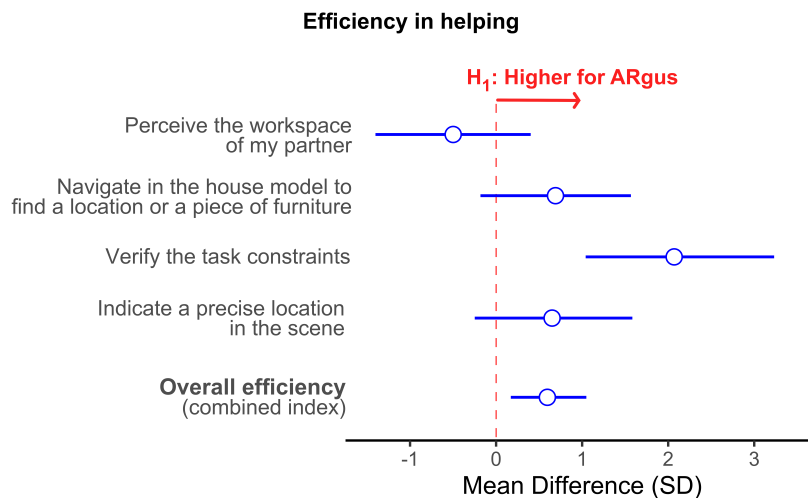


Figure 4.14: Comparing the importance of verbal communication for the two user interface configurations ($N = 12$). We use again Bayesian ordinal (cumulative probit) models [25]. The bars in the graph represent 95% credible intervals of mean differences over a latent continuous variable and can be treated as estimates of standardized effect sizes.

Our transcript analysis provides additional information about how participants verbally communicated instructions. Figure 4.15 summa-

rizes our results. Overall, the ARGUS user interface reduced the number of words that belonged to instructions by 151.8, 95% CI [25.7, 278.0], $t(11) = 2.65$, $p = .023$ (see Hypothesis H_3). To put this number in perspective, participants pronounced on average 834.5 words with the HEADSET VIEW, where 435.6 of these words were instructions. We observe that clear differences between conditions only concern instructions that ask the experimenter to move around the model. Surprisingly, there is no clear difference in the number of words used by participants to guide the experimenter on how to identify, reach, and manipulate (e.g., translate or rotate) objects. A possible explanation of this result is the fact that five participants did not at all use the stick (see Figure 4.12) and relied on verbal instructions for these subtasks. Indeed, a post hoc analysis shows a strong correlation between the use of the stick (binary variable) and the difference of words used for these subtasks (*Point-biserial correlation* = .79, 95% CI [.40, .94]). Seven participants who used the stick pronounced 156.9 fewer words (95% CI [41.8, 271.9]) with ARGUS when they provided instructions for these subtasks. This result, however, must be treated with caution because uncontrolled ordering effects may exaggerate the difference.

4.4.8 Discussion

Overall, our results confirm that remote desktop collaborators can benefit from the multiple views of ARGUS, since each view is best adapted to a different aspect of the task. The VIRTUAL VIEW makes navigation in the virtual model easier and independent of the position and visual focus of the local AR user. The EXTERNAL VIEW provides a static overview of the workspace, showing both virtual and physical objects. Finally, the HEADSET VIEW allows remote users to directly observe the view and actions of their local partner and provide direct instructions. Our participants demonstrated various strategies on how to combine these views with the tools of ARGUS. Given previous results [67, 223], we expected a more extensive use of the EXTERNAL VIEW. However, using all three views can be complex, increasing cognitive costs. So many participants judged that the VIRTUAL VIEW and the HEADSET VIEW were enough for completing the task. Nevertheless, mastering all combinations of views and previews, as well as developing strategies to use them effectively in various steps of the collaboration, may require a long learning process that we did not assess in our studies. Finding a good viewpoint for an external camera also remains a problem. A solution may be to reposition the external camera on the fly depending on the collaborative situation, as explored in Giusti et al. [80]. The nature of the task may also explain why most participants largely relied on the VIRTUAL VIEW to complete the task. It is reasonable to expect that if key objects and landmarks in the scene were mostly physical rather than virtual, the VIRTUAL VIEW

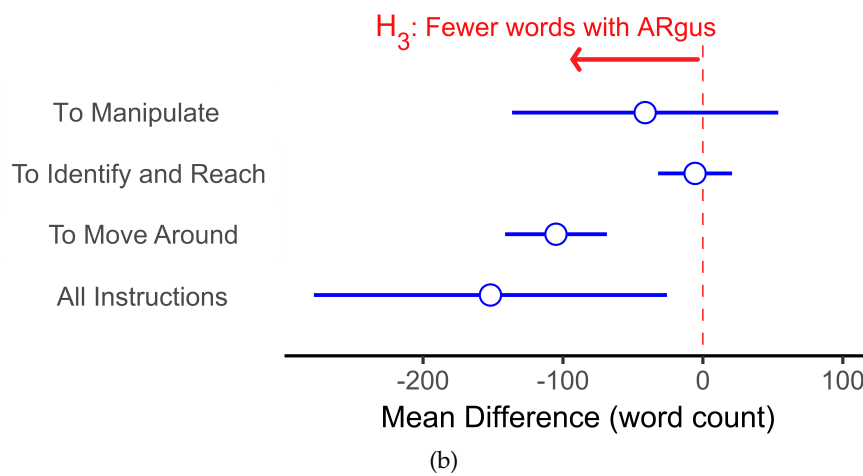
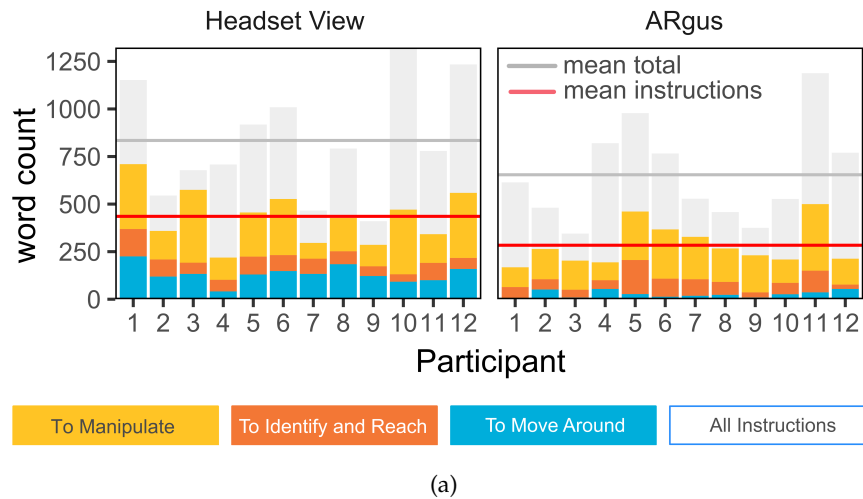


Figure 4.15: **Results of transcript analysis.** We compare the number of words pronounced by the 12 participants to provide instructions. The grey boxes at the left show the total number of words with non-instructions. The error bars at the right represent 95% confidence intervals derived from the t -distribution.

might be less appropriate, while the two other views might be more frequently used. Clearly, there are trade-offs in the choice of each view that largely depend on where the task falls in the continuum between virtual and physical.

The results support our three hypotheses. Participants perceived on average that ARGUS was more efficient than the control HEADSET VIEW condition (H_1) and lessened the importance of verbal communication (H_2). We also found that ARGUS reduced the average number of words of remote instructions (H_3), which corroborates previous evidence [227] that increased view independence reduces the prevalence of verbal instructions.

We acknowledge that our experimental method and setup present several limitations. The experimenter took the role of the local collabo-

rator in all experimental sessions, which inevitably limits the external validity of our results. The variable quality of the internet connection and the limited resolution of the HoloLens frontal camera may have had an effect as well. Furthermore, we studied one only part of the bilateral collaboration, neglecting how the local AR user perceives and interprets instructions given by the remote collaborator through multiple complementary views. Future user studies should thus examine the collaboration strategies (verbal communication, physical navigation, and gestural interaction) of local users, and their need for awareness of remote user actions. Finally, we are interested in enriching ARgus' pointing, annotation, and hybrid navigation tools and evaluate their collaboration effectiveness with more specialized experimental tasks.

CONCLUSION AND FUTURE DIRECTIONS

My dissertation explored new ways of enhancing collaboration through the use of Augmented Reality (AR) technologies. The aim was to expand the boundaries of interaction design beyond what is achievable in the physical world, with a specific focus on collaborative design tasks. I studied two research directions based on the spatial proximity of collaborators. The first direction investigated co-located collaboration, while the second looked at remote collaboration. In each case, I identified the primary limitations in past research and designed new interactive systems to address them.

In the context of co-located collaboration, design tasks raise several issues about visualization and privacy. In particular, co-located users experience occlusion and space occupancy conflicts when they simultaneously modify the same virtual object or create new content around the same location. Their previous creations can also compete for space with their current design explorations. AR enables users to see and interact with multiple virtual spaces while staying at the same physical space. I argue that bringing this insight into the context of collaborative design can help us address the aforementioned issues.

In order to achieve this objective, I created a conceptual framework for collaborative design in AR that enables multiple versions of virtual artifacts. These versions correspond to different states of the collaborators' virtual space and are reified as Version Objects. These objects can be manipulated in space, arranged in collections, and compared with each other using preview tools or 3D portals. The conceptual framework also allows collaborators to desynchronize or synchronize their virtual spaces. They can thus create and perceive different augmentations while evolving in the same physical space. All these concepts enable them to visualize overlapping versions of a 3D model and to control the visibility of their creations and their workspace. I proposed an implementation of this framework and demonstrated its use through a scenario in which two AR headset users generate ideas by sketching in 3D on a sewing mannequin.

In the context of remote collaboration, tasks that involve physical and virtual content raise concerns about the workspace awareness of the remote users. They rely either on a workspace reconstruction or on a video-based technique. The first one is limited by its technical complexity and its dependence on the reconstruction quality, while the second one is limited by the fixed viewpoint of the camera. I argue that combining pointing/annotating tools with multiple views that constitute a trade-off among workspace representation, viewpoint,

and navigation control, may address these limitations. To explore this problem, I targeted a scenario where a remote desktop user wants to collaborate with an AR headset user by interacting with both physical and virtual content. I first conducted a user study that compares three view representations for the remote desktop user: (i) a first-person view showing augmented video from the AR headset, (ii) a third-person view displaying augmented video from an external camera, and (iii) a virtual view providing a free viewpoint on a fully virtual representation. Structured as two independent sub-studies with 12 participants each, this study confirmed that each view presents different benefits, targeting different aspects of the collaboration. Based on these insights, I developed ARgus, a multi-view collaboration system that provides tools for effectively switching among views, virtually navigating in the remote AR workspace, pointing, and annotating 3D models. I then ran a second user study to evaluate how 12 remote participants used ARgus to instruct a local user wearing an AR headset to arrange physical furniture miniatures in a virtual house model. I observed that participants frequently switched among views or concurrently used them through ARgus preview functionality. These results also suggest that the added flexibility of ARgus multi-view interface allows remote users to verify spatial constraints more efficiently and reduces their reliance on verbal instructions.

Below, I discuss limitation of this work and identify promising research directions for the future.

5.1 LIMITATIONS

In the context of collaborative design, pushing the limits of interaction beyond what is possible in the physical world can enhance collaboration, as presented previously. However, it is important to acknowledge the limitations that arise from breaking away from the constraints and rules of the physical world. In particular, it can destabilize users who are accustomed to such constraints and rules. This section discusses the limitations that can arise when going beyond the unique instance of space that is available in the physical world and a single user viewpoint.

5.1.1 *Beyond a unique space*

Breaking the constraint of a unique 3D space and allowing multiple parallel versions of objects to exist can first create limitations in the way users interact with the virtual content. In chapter 3, Version Objects were proposed as a means of creating and managing alternatives of augmentations. While traditional 3D model manipulation can be used to move Version Objects in space and organize them, the presence of numerous alternatives may cause cognitive overload by providing

a large amount of information and hindering users' selection process. Additionally, these objects may obstruct virtual content if placed between users' viewpoints and objects of interest. Reifying groups of objects or adapting the visibility of 3D objects could be further investigated to address these limitations.

Likewise, many alternatives may generate difficulties in identifying each of them. In chapter 3, I introduce a technique to retrieve alternatives that do or do not share selected 3D content. A high number of retrieved alternatives might raise the same issues as the ones described in the previous paragraph. Other techniques [79, 101] could be used to reduce these effects.

Furthermore, Version Objects only consider alternatives to virtual content. This means that any changes made to the physical environment are visible to all collaborators, which may not always be desirable. To address this issue, virtuality can provide partial support. The first approach is to switch to a virtual representation of the physical object and make modifications effective on the virtual 3D model, as is done with proxies [62, 168]. Another approach is to rely on the ability of Augmented Reality to hide physical content [87]. In video see-through Augmented Reality, object removal and inpainting techniques can make physical objects partially or fully disappear from users' view, letting them experience a *Diminished Reality* [91, 116]. Co-located users in *Diminished Reality* could then perceive different appearances of the same physical object and thus consider various alternatives simultaneously. However, physical objects or part of physical objects remain tangible even if they are not visible. Further investigation should be done to address this limitation.

Allowing for multiple virtual spaces at a time can also hinder collaboration among users. "One "play it safe" design strategy is to strictly adhere to a physical world metaphor so that our everyday intuitions can be carried over into the medium supplied by the technology" [212]. This statement from Randall B. Smith aims to support real-time interaction by considering all the "subtleties of human communication" in a collaborative scenario. The framework presented in the chapter 3 does not apply this "What You See Is What I Think You See" principle to provide new interaction for collaboration, such as privacy or parallel creation. This disparity with the physical world may damage communication and consequently understanding among collaborators.

Having desynchronized virtual spaces among collaborators can also lead to conflicts in physical space occupation. Empty spaces in a user's virtual space may not be empty in their collaborators' virtual spaces, causing users to occupy the same space and potentially interfere with their collaborator's work. This effect may be especially pronounced with small-scaled augmentations. While such limits could be addressed with feedback, they still require trade-offs between shared information and workspace privacy.

5.1.2 *Beyond a single viewpoint*

ARgus enables users to seamlessly switch among different viewpoints and different representations, and this can affect the way users perceive space. For instance, when users switch from an augmented-video view to a virtual view, they can no longer perceive the physical space. This can be problematic for several reasons. First, users lose the shared context that comes from being able to see dynamic changes in the physical environment. Then, this loss of context also hinders users' ability to point and annotate using physical objects as reference points. Furthermore, when switching from the third-person view to another view, users lose the facial expression and body language related to the body parts that are not tracked. These limitations hamper the workspace awareness and may have a negative impact on collaboration.

One way to address these limitations is by using a virtual reconstruction of the physical world for the virtual view, as presented in chapter 2. However, as argued in the chapter 4, the quality of reconstruction and the complexity of such techniques can hinder the collaboration. However, reconstruction of an entire space may not be necessary as only the features related to the practical tasks may be needed [45]. Partial reconstruction may then be considered to address this situation.

5.2 FUTURE WORK

One possible direction for further research is to consider multi-user interactions in AR, building on the interactions described in previous chapters but involving more than two collaborators. In Chapter 4, for instance, multiple users could interact in the same virtual workspace using either AR headsets or desktop computers. However, one of the key challenges here is how to represent each user in the virtual space, as currently only the AR user's head and hands are visible. To avoid confusion, annotations and pointing sticks should be associated with their respective creators. Additionally, future work could investigate how to resolve any conflicts that might arise when multiple users interact within the same AR space.

In Chapter 3, an important issue is how to determine which collaborators are synchronized with each other. Further research could explore how users share both their virtual and physical spaces when not all are synchronized, and how social interactions between synchronized and desynchronized collaborators in the same physical space might be affected.

Both the ARgus and framework developed in this dissertation could be integrated to enable the creation and management of 3D sketches in remote AR collaboration, using Version Objects. This offers new

design opportunities to expand content creation on traditional desktop CAD software in an augmented remote workspace.

Beyond these specific directions, this dissertation also presents interactions that depart from traditional physical metaphors. Further research could explore additional interactions based on this approach, for example, by investigating geometric transformations of the virtual space in which AR collaborators operate. While previous work has explored this idea for single-user AR environments, there is potential for this to be extended to multi-user AR as well. Additionally, research could explore non-physical metaphors as a design strategy for creating new interactions in collaborative AR, as this area has received less attention to date. By thinking about metaphors and their limitations, designers may be able to create new and innovative interactions that can enhance collaborative AR experiences.

PUBLICATIONS

- [1] Arthur Fages, Cédric Fleury, and Theophanis Tsandilas. “AR-gus : système multi-vues pour collaborer à distance avec un utilisateur en réalité augmentée.” In: *IHM '22 - 33ème Conférence Francophone sur l'Interaction Homme-Machine. Demonstration*. Namur, Belgium, Apr. 2022. URL: <https://hal.science/hal-03762816>.
- [2] 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*. CSCW2 549 (Nov. 2022), 27 pages. DOI: [10.1145/3555607](https://doi.org/10.1145/3555607). URL: <https://hal.science/hal-03762803>.
- [3] Arthur Fages, Cédric Fleury, and Theophanis Tsandilas. “Conception collaborative au travers de versions parallèles en Réalité Augmentée.” In: *Proceedings of the 34th Conference on l'Interaction Humain-Machine. IHM '23*. Troyes, France: Association for Computing Machinery, 2023. ISBN: 978-1-4503-9824-4/23/04. DOI: [10.1145/3583961.3583978](https://doi.org/10.1145/3583961.3583978).

BIBLIOGRAPHY

- [1] Matt Adcock, Stuart Anderson, and Bruce Thomas. "Remote-Fusion: Real Time Depth Camera Fusion for Remote Collaboration on Physical Tasks." In: *Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*. VRCAI '13. Hong Kong, Hong Kong: Association for Computing Machinery, 2013, 235–242. ISBN: 9781450325905. DOI: [10.1145/2534329.2534331](https://doi.org/10.1145/2534329.2534331). URL: <https://doi.org/10.1145/2534329.2534331>.
- [2] Anand Agarawala and Ravin Balakrishnan. "Keepin' It Real: Pushing the Desktop Metaphor with Physics, Piles and the Pen." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '06. Montréal, Québec, Canada: Association for Computing Machinery, 2006, 1283–1292. ISBN: 1595933727. DOI: [10.1145/1124772.1124965](https://doi.org/10.1145/1124772.1124965). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/1124772.1124965>.
- [3] 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: *Proc. ACM Hum.-Comput. Interact.* 5.CSCW2 (2021). DOI: [10.1145/3476054](https://doi.org/10.1145/3476054). URL: <https://doi.org/10.1145/3476054>.
- [4] Canan Akoglu. "The Relationship between Industrial Design and Interaction Design in Product Development Activities." In: *CHI '12 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '12. Austin, Texas, USA: Association for Computing Machinery, 2012, 769–776. ISBN: 9781450310161. DOI: [10.1145/2212776.2212850](https://doi.org/10.1145/2212776.2212850). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/2212776.2212850>.
- [5] Irwin Altman. "The environment and social behavior: privacy, personal space, territory, and crowding." In: (1975).
- [6] J.L Alty, R.P Knott, B Anderson, and M Smyth. "A framework for engineering metaphor at the user interface." In: *Interacting with Computers* 13.2 (2000), pp. 301–322. ISSN: 0953-5438. DOI: [https://doi.org/10.1016/S0953-5438\(00\)00047-3](https://doi.org/10.1016/S0953-5438(00)00047-3). URL: <https://www.sciencedirect.com/science/article/pii/S095354380000473>.
- [7] Lynda M. Applegate. "Technology support for cooperative work: A framework for studying introduction and assimilation in organizations." In: *Journal of Organizational Computing* 1.1 (1991), pp. 11–39. DOI: [10.1080/10919399109540148](https://doi.org/10.1080/10919399109540148). eprint:

<https://doi.org/10.1080/10919399109540148>. URL: <https://doi.org/10.1080/10919399109540148>.

- [8] 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 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Honolulu, HI, USA: Association for Computing Machinery, 2020, 1–13. ISBN: 9781450367080. DOI: [10.1145/3313831.3376550](https://doi.org/10.1145/3313831.3376550). URL: <https://doi.org/10.1145/3313831.3376550>.
- [9] Istvan Barakonyi, Tamer Fahmy, and Dieter Schmalstieg. "Remote Collaboration Using Augmented Reality Videoconferencing." In: *Proceedings of Graphics Interface 2004*. GI '04. London, Ontario, Canada: Canadian Human-Computer Communications Society, 2004, 89–96. ISBN: 1568812272.
- [10] M. Billinghurst, S. Weghorst, and T. Furness. "Shared space: An augmented reality approach for computer supported collaborative work." In: *Virtual Reality 3.1* (Mar. 1998), pp. 25–36. ISSN: 1434-9957. DOI: [10.1007/BF01409795](https://doi.org/10.1007/BF01409795). URL: <https://doi.org/10.1007/BF01409795>.
- [11] Mark Billinghurst and Hirokazu Kato. "Collaborative mixed reality." In: *Proceedings of the first international symposium on mixed reality*. 1999, pp. 261–284.
- [12] Mark Billinghurst, Hirokazu Kato, Kiyoshi Kiyokawa, Daniel Belcher, and Ivan Poupyrev. "Experiments with Face-To-Face Collaborative AR Interfaces." In: *Virtual Reality 6* (Oct. 2002), pp. 107–121. DOI: [10.1007/s100550200012](https://doi.org/10.1007/s100550200012).
- [13] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. "The MagicBook: a transitional AR interface." In: *Computers & Graphics 25.5* (2001). Mixed realities - beyond conventions, pp. 745–753. ISSN: 0097-8493. DOI: [https://doi.org/10.1016/S0097-8493\(01\)00117-0](https://doi.org/10.1016/S0097-8493(01)00117-0). URL: <https://www.sciencedirect.com/science/article/pii/S0097849301001170>.
- [14] Mark Billinghurst, Ivan Poupyrev, Hirokazu Kato, and Richard May. "Mixing Realities in Shared Space: An Augmented Reality Interface for Collaborative Computing." In: vol. 3. Jan. 2000, pp. 1641–1644. DOI: [10.1109/ICME.2000.871085](https://doi.org/10.1109/ICME.2000.871085).
- [15] Nickolas Bloom. *Working from Home and the Future of U.S. Economic Growth under COVID*. <https://www.youtube.com/watch?v=jtdFIZx3hyk>. 2020.
- [16] Charlotte Boddien, Jill Heitmann, Florian Hermuth, Dawid Lokiec, Carlos Tan, Laura Wölbeling, Thomas Jung, and Johann Habakuk Israel. "SketchTab3d: A Hybrid Sketch Library Using Tablets and Immersive 3D Environments." In: *Proceedings of the 2017 ACM Symposium on Document Engineering*. DocEng '17.

- Valletta, Malta: Association for Computing Machinery, 2017, 101–104. ISBN: 9781450346894. DOI: [10.1145/3103010.3121029](https://doi.org/10.1145/3103010.3121029). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3103010.3121029>.
- [17] Kenneth R Boff. “The tower of Babel revisited: On crossdisciplinary chokepoints in system design.” In: *System design: Behavioral perspectives on designers, tools, and organizations* (1987), pp. 83–96.
- [18] Adrien Bousseau, Theophanis Tsandilas, Lora Oehlberg, and Wendy E. Mackay. “How Novices Sketch and Prototype Hand-Fabricated Objects.” In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI ’16. San Jose, California, USA: Association for Computing Machinery, 2016, 397–408. ISBN: 9781450333627. DOI: [10.1145/2858036.2858159](https://doi.org/10.1145/2858036.2858159). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/2858036.2858159>.
- [19] Tobias Brockmann, Nina Krüger, Stefan Stieglitz, and Immo Bohlson. “A Framework for Collaborative Augmented Reality Applications.” In: vol. 1. Aug. 2013.
- [20] Wolfgang Broll, Irma Lindt, Jan Ohlenburg, Michael Wittkamper, Chunrong Yuan, Thomas Novotny, Ava Fatah gen. Schieck, Chiron Mottram, and A. Strothman. “ARTHUR: A Collaborative Augmented Environment for Architectural Design and Urban Planning.” In: *Journal of Virtual Reality and Broadcasting* 1 (Dec. 2004), pp. 1–10.
- [21] Wolfgang Broll, Eckhard Meier, and Thomas Schardt. “The Virtual Round Table - a Collaborative Augmented Multi-User Environment.” In: *Proceedings of the Third International Conference on Collaborative Virtual Environments*. CVE ’00. San Francisco, California, USA: Association for Computing Machinery, 2000, 39–45. ISBN: 1581133030. DOI: [10.1145/351006.351011](https://doi.org/10.1145/351006.351011). URL: <https://doi.org/10.1145/351006.351011>.
- [22] Gordon Brown and Michael Prilla. “Evaluating Pointing Modes and Frames of Reference for Remotely Supporting an Augmented Reality User in a Collaborative (Virtual) Environment: Evaluation within the Scope of a Remote Consultation Session.” In: *Proceedings of Mensch Und Computer 2019*. MuC’19. Hamburg, Germany: Association for Computing Machinery, 2019, 713–717. ISBN: 9781450371988. DOI: [10.1145/3340764.3344896](https://doi.org/10.1145/3340764.3344896). URL: <https://doi.org/10.1145/3340764.3344896>.
- [23] Jean-Marie Burkhardt, Françoise Détienne, Linda Moutsingua-Mpaga, Laurence Perron, Stéphane Safin, and Pierre Leclercq. “Multimodal Collaborative Activity among Architectural Designers Using an Augmented Desktop at Distance or in Collocation.” In: *Proceedings of the 15th European Conference on Cognitive*

- Ergonomics: The Ergonomics of Cool Interaction*. ECCE '08. Funchal, Portugal: Association for Computing Machinery, 2008. ISBN: 9781605583990. DOI: [10.1145/1473018.1473049](https://doi.org/10.1145/1473018.1473049). URL: <https://doi.org/10.1145/1473018.1473049>.
- [24] William A. S. Buxton. "Telepresence: Integrating Shared Task and Person Spaces." In: *Proceedings of the Conference on Graphics Interface '92*. Vancouver, British Columbia, Canada: Morgan Kaufmann Publishers Inc., 1992, 123–129. ISBN: 0969533810.
- [25] Paul-Christian Bürkner and Matti Vuorre. "Ordinal Regression Models in Psychology: A Tutorial." In: *Advances in Methods and Practices in Psychological Science* 2.1 (2019), pp. 77–101. DOI: [10.1177/2515245918823199](https://doi.org/10.1177/2515245918823199). eprint: <https://doi.org/10.1177/2515245918823199>. URL: <https://doi.org/10.1177/2515245918823199>.
- [26] Yuanzhi Cao, Tianyi Wang, Xun Qian, Pawan S. Rao, Manav Wadhawan, Ke Huo, and Karthik Ramani. "GhostAR: A Time-Space Editor for Embodied Authoring of Human-Robot Collaborative Task with Augmented Reality." In: *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. UIST '19. New Orleans, LA, USA: Association for Computing Machinery, 2019, 521–534. ISBN: 9781450368162. DOI: [10.1145/3332165.3347902](https://doi.org/10.1145/3332165.3347902). URL: <https://doi.org/10.1145/3332165.3347902>.
- [27] Géry Casiez, Nicolas Roussel, and Daniel Vogel. "1 € Filter: A Simple Speed-Based Low-Pass Filter for Noisy Input in Interactive Systems." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '12. Austin, Texas, USA: Association for Computing Machinery, 2012, 2527–2530. ISBN: 9781450310154. DOI: [10.1145/2207676.2208639](https://doi.org/10.1145/2207676.2208639). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/2207676.2208639>.
- [28] C.D. Cera, W.C. Regli, I. Braude, Y. Shapirstein, and C.V. Foster. "A collaborative 3D environment for authoring design semantics." In: *IEEE Computer Graphics and Applications* 22.3 (2002), pp. 43–55. ISSN: 1558-1756. DOI: [10.1109/MCG.2002.999787](https://doi.org/10.1109/MCG.2002.999787).
- [29] Yoonjeong Cha, Sungu Nam, Mun Yong Yi, Jaeseung Jeong, and Woontack Woo. "Augmented Collaboration in Shared Space Design with Shared Attention and Manipulation." In: *The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings*. UIST '18 Adjunct. Berlin, Germany: Association for Computing Machinery, 2018, 13–15. ISBN: 9781450359498. DOI: [10.1145/3266037.3266086](https://doi.org/10.1145/3266037.3266086). URL: <https://doi.org/10.1145/3266037.3266086>.

- [30] Senthil Chandrasegaran, Devarajan Ramanujan, and Niklas Elmqvist. "How Do Sketching and Non-Sketching Actions Convey Design Intent?" In: *Proceedings of the 2018 Designing Interactive Systems Conference*. DIS '18. Hong Kong, China: Association for Computing Machinery, 2018, 373–385. ISBN: 9781450351980. DOI: [10.1145/3196709.3196723](https://doi.org/10.1145/3196709.3196723). URL: <https://doi.org/10.1145/3196709.3196723>.
- [31] T. Chassin, J. Ingensand, M. Lotfian, O. Ertz, and F. Joerin. "Challenges in creating a 3D participatory platform for urban development." In: *Advances in Cartography and GIScience of the ICA 1* (2019), p. 3. DOI: [10.5194/ica-adv-1-3-2019](https://doi.org/10.5194/ica-adv-1-3-2019). URL: <https://www.adv-cartogr-giscience-int-cartogr-assoc.net/1/3/2019/>.
- [32] Jeff Chastine, Kristine Nagel, Ying Zhu, and Mary Hudachek-Buswell. "Studies on the Effectiveness of Virtual Pointers in Collaborative Augmented Reality." In: *2008 IEEE Symposium on 3D User Interfaces*. 2008, pp. 117–124. DOI: [10.1109/3DUI.2008.4476601](https://doi.org/10.1109/3DUI.2008.4476601).
- [33] Jeffrey W. Chastine, Kristine Nagel, Ying Zhu, and Luca Yearso-vich. "Understanding the Design Space of Referencing in Collaborative Augmented Reality Environments." In: *Proceedings of Graphics Interface 2007*. GI '07. Montreal, Canada: Association for Computing Machinery, 2007, 207–214. ISBN: 9781568813370. DOI: [10.1145/1268517.1268552](https://doi.org/10.1145/1268517.1268552). URL: <https://doi.org/10.1145/1268517.1268552>.
- [34] Chih-Fan Chen, Mark Bolas, and Evan Suma. "Real-Time 3D Rendering Using Depth-Based Geometry Reconstruction and View-Dependent Texture Mapping." In: *ACM SIGGRAPH 2016 Posters*. SIGGRAPH '16. Anaheim, California: Association for Computing Machinery, 2016. ISBN: 9781450343718. DOI: [10.1145/2945078.2945162](https://doi.org/10.1145/2945078.2945162). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/2945078.2945162>.
- [35] Shaoyu Chen et al. "UrbanRama: Navigating Cities in Virtual Reality." In: *IEEE Transactions on Visualization and Computer Graphics* 28.12 (2022), pp. 4685–4699. ISSN: 1941-0506. DOI: [10.1109/TVCG.2021.3099012](https://doi.org/10.1109/TVCG.2021.3099012).
- [36] Morgan Le Chenechal, Thierry Duval, Valerie Gouranton, Jerome Royan, and Bruno Arnaldi. "Vishnu: virtual immersive support for HelpiNg users an interaction paradigm for collaborative remote guiding in mixed reality." In: *2016 IEEE Third VR International Workshop on Collaborative Virtual Environments (3DCVE)*. 2016, pp. 9–12. DOI: [10.1109/3DCVE.2016.7563559](https://doi.org/10.1109/3DCVE.2016.7563559).

- [37] Alexis Clay, Sébastien Bottecchia, Régis Mollard, and Marion Wolff. “Évaluation d’un prototype (système TAC) de communication en réalité augmentée entre un opérateur et un expert pour l’aide à la maintenance aéronautique.” In: July 2016.
- [38] Andy Cockburn, Carl Gutwin, and Alan Dix. “HARK No More: On the Preregistration of CHI Experiments.” In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2018, 1–12. ISBN: 9781450356206. URL: <https://doi.org/10.1145/3173574.3173715>.
- [39] Verina Cristie and Sam Conrad Joyce. “Versioning for parametric design exploration process.” In: *Automation in Construction* 129 (2021), p. 103802. ISSN: 0926-5805. DOI: <https://doi.org/10.1016/j.autcon.2021.103802>. URL: <https://www.sciencedirect.com/science/article/pii/S0926580521002533>.
- [40] Christian Heide Damm, Klaus Marius Hansen, and Michael Thomsen. “Tool Support for Cooperative Object-Oriented Design: Gesture Based Modelling on an Electronic Whiteboard.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’00. The Hague, The Netherlands: Association for Computing Machinery, 2000, 518–525. ISBN: 1581132166. DOI: [10.1145/332040.332488](https://doi.org/10.1145/332040.332488). URL: <https://doi.org/10.1145/332040.332488>.
- [41] Kurtis Danyluk, Barrett Ens, Bernhard Jenny, and Wesley Willett. “A Design Space Exploration of Worlds in Miniature.” In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI ’21. Yokohama, Japan: Association for Computing Machinery, 2021. ISBN: 9781450380966. DOI: [10.1145/3411764.3445098](https://doi.org/10.1145/3411764.3445098). URL: <https://doi.org/10.1145/3411764.3445098>.
- [42] Françoise Darses, Françoise Détienne, P. Falzon, and Willemien Visser. *COMET. A Method for Analysing Collective Design Processes*. Research Report RR-4258. Projet EIFFEL. INRIA, 2001. URL: <https://hal.inria.fr/inria-00072330>.
- [43] Fabian Lorenzo Dayrit, Yuta Nakashima, Tomokazu Sato, and Naokazu Yokoya. “Free-viewpoint AR human-motion reenactment based on a single RGB-D video stream.” In: *2014 IEEE International Conference on Multimedia and Expo (ICME)*. 2014, pp. 1–6. DOI: [10.1109/ICME.2014.6890243](https://doi.org/10.1109/ICME.2014.6890243).
- [44] Edward De Bono. “The Use of Lateral Thinking.” In: *London: Cape* (1967).

- [45] Giorgio De Michelis, Flavio De Paoli, Costanza Pluchinotta, and Marco Susani. "Weakly Augmented Reality: Observing and Designing the Work-Place of Creative Designers." In: *Proceedings of DARE 2000 on Designing Augmented Reality Environments*. DARE '00. Elsinore, Denmark: Association for Computing Machinery, 2000, 81–91. ISBN: 9781450373265. DOI: [10.1145/354666.354675](https://doi.org/10.1145/354666.354675). URL: <https://doi.org/10.1145/354666.354675>.
- [46] Françoise Détienne. "Collaborative design: Managing task interdependencies and multiple perspectives." In: *ArXiv abs/cs/0611151* (2006).
- [47] Arindam Dey, Graeme Jarvis, Christian Sandor, Ariawan Wibowo, and Ville-Veikko Mattila. "An Evaluation of Augmented Reality X-Ray Vision for Outdoor Navigation." In: Nov. 2011, pp. 28–32.
- [48] Veronika Domova, Elina Vartiainen, and Marcus Englund. "Designing a Remote Video Collaboration System for Industrial Settings." In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*. ITS '14. Dresden, Germany: Association for Computing Machinery, 2014, 229–238. ISBN: 9781450325875. DOI: [10.1145/2669485.2669517](https://doi.org/10.1145/2669485.2669517). URL: <https://doi.org/10.1145/2669485.2669517>.
- [49] Tomas Dorta, Gokce Kinayoglu, and Michael Hoffmann. "Hyve-3D: A New Embodied Interface for Immersive Collaborative 3D Sketching." In: *ACM SIGGRAPH 2014 Studio*. SIGGRAPH '14. Vancouver, Canada: Association for Computing Machinery, 2014. ISBN: 9781450329774. DOI: [10.1145/2619195.2656325](https://doi-org.ins2i.bib.cnrs.fr/10.1145/2619195.2656325). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/2619195.2656325>.
- [50] Mingsong Dou et al. "Fusion4D: Real-Time Performance Capture of Challenging Scenes." In: *ACM Trans. Graph.* 35.4 (2016). ISSN: 0730-0301. DOI: [10.1145/2897824.2925969](https://doi.org/10.1145/2897824.2925969). URL: <https://doi.org/10.1145/2897824.2925969>.
- [51] Steven Dow, Julie Fortuna, Dan Schwartz, Beth Altringer, Daniel Schwartz, and Scott Klemmer. "Prototyping Dynamics: Sharing Multiple Designs Improves Exploration, Group Rapport, and Results." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '11. Vancouver, BC, Canada: Association for Computing Machinery, 2011, 2807–2816. ISBN: 9781450302289. DOI: [10.1145/1978942.1979359](https://doi.org/10.1145/1978942.1979359). URL: <https://doi.org/10.1145/1978942.1979359>.
- [52] John J. Dudley, Hendrik Schuff, and Per Ola Kristensson. "Bare-Handed 3D Drawing in Augmented Reality." In: *Proceedings of the 2018 Designing Interactive Systems Conference*. DIS '18. Hong Kong, China: Association for Computing Machinery, 2018, 241–252. ISBN: 9781450351980. DOI: [10.1145/3196709.3196737](https://doi.org/10.1145/3196709.3196737).

- URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3196709.3196737>.
- [53] Claudia Eckert and Martin Stacey. "Sources of Inspiration: A Language of Design." In: *Design Studies* 21 (Sept. 2000). DOI: [10.1016/S0142-694X\(00\)00022-3](https://doi.org/10.1016/S0142-694X(00)00022-3).
- [54] Jörg Edelmann, Philipp Mock, Andreas Schilling, and Peter Gerjets. "Preserving Non-Verbal Features of Face-to-Face Communication for Remote Collaboration." In: *Proceedings of the 10th International Conference on Cooperative Design, Visualization, and Engineering - Volume 8091*. Berlin, Heidelberg: Springer-Verlag, 2013, 27–34. ISBN: 9783642408397. DOI: [10.1007/978-3-642-40840-3_4](https://doi.org/10.1007/978-3-642-40840-3_4). URL: https://doi.org/10.1007/978-3-642-40840-3_4.
- [55] K. Eissen and R. Steur. *Sketching: The Basics*. BIS, 2011. ISBN: 9789063692537. URL: <https://books.google.fr/books?id=pigvnwEACAAJ>.
- [56] Niklas Elmqvist. "BalloonProbe: Reducing Occlusion in 3D Using Interactive Space Distortion." In: *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. VRST '05. Monterey, CA, USA: Association for Computing Machinery, 2005, 134–137. ISBN: 1595930981. DOI: [10.1145/1101616.1101643](https://doi.org/10.1145/1101616.1101643). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/1101616.1101643>.
- [57] Niklas Elmqvist, Ulf Assarsson, and Philippas Tsigas. "Dynamic Transparency for 3D Visualization: Design and Evaluation." In: *International Journal of Virtual Reality - IJVR* 8 (Jan. 2009). DOI: [10.20870/IJVR.2009.8.1.2715](https://doi.org/10.20870/IJVR.2009.8.1.2715).
- [58] Arthur Fages, Cédric Fleury, and Theophanis Tsandilas. "Argus : système multi-vues pour collaborer à distance avec un utilisateur en réalité augmentée." In: *IHM '22 - 33ème Conférence Francophone sur l'Interaction Homme-Machine*. Demonstration. Namur, Belgium, Apr. 2022. URL: <https://hal.science/hal-03762816>.
- [59] 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*. CSCW2 549 (Nov. 2022), 27 pages. DOI: [10.1145/3555607](https://doi.org/10.1145/3555607). URL: <https://hal.science/hal-03762803>.
- [60] Arthur Fages, Cédric Fleury, and Theophanis Tsandilas. "Conception collaborative au travers de versions parallèles en Réalité Augmentée." In: *Proceedings of the 34th Conference on l'Interaction Humain-Machine*. IHM '23. Troyes, France: Association for Com-

- puting Machinery, 2023. ISBN: 978-1-4503-9824-4/23/04. DOI: [10.1145/3583961.3583978](https://doi.org/10.1145/3583961.3583978).
- [61] Allen Fairchild, Simon Champion, Arturo García, Robin Wolff, T. Fernando, and David Roberts. “A Mixed Reality Telepresence System for Collaborative Space Operation.” In: *IEEE Transactions on Circuits and Systems for Video Technology* 27 (Jan. 2016), pp. 1–1. DOI: [10.1109/TCSVT.2016.2580425](https://doi.org/10.1109/TCSVT.2016.2580425).
- [62] Martin Feick, Anthony Tang, and Scott Bateman. “Mixed-Reality for Object-Focused Remote Collaboration.” In: *The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings*. UIST ’18 Adjunct. Berlin, Germany: Association for Computing Machinery, 2018, 63–65. ISBN: 9781450359498. DOI: [10.1145/3266037.3266102](https://doi.org/10.1145/3266037.3266102). URL: <https://doi.org/10.1145/3266037.3266102>.
- [63] M. Fiorentino, R. de Amicis, G. Monno, and A. Stork. “Spacedesign: a mixed reality workspace for aesthetic industrial design.” In: *Proceedings. International Symposium on Mixed and Augmented Reality*. 2002, pp. 86–318. DOI: [10.1109/ISMAR.2002.1115077](https://doi.org/10.1109/ISMAR.2002.1115077).
- [64] 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: *ICAT 2010 (20th International Conference on Artificial Reality and Telexistence)* (Dec. 2010).
- [65] Mike Fraser, Steve Benford, Jon Hindmarsh, and Christian Heath. “Supporting Awareness and Interaction through Collaborative Virtual Interfaces.” In: *Proceedings of the 12th Annual ACM Symposium on User Interface Software and Technology*. UIST ’99. Asheville, North Carolina, USA: Association for Computing Machinery, 1999, 27–36. ISBN: 1581130759. DOI: [10.1145/320719.322580](https://doi.org/10.1145/320719.322580). URL: <https://doi.org/10.1145/320719.322580>.
- [66] Henry Fuchs, Gary Bishop, Kevin Arthur, Leonard Mcmillan, Ruzena Bajcsy, Sangwook Lee, Hany Farid, and Takeo Kanade. “Virtual Space Teleconferencing using a Sea of Cameras.” In: *Proceeding of First International Conference on Medical Robotics and Computer Assisted Surgery, Pittsburgh 2* (Sept. 1994).
- [67] Susan R. Fussell, Leslie D. Setlock, and Robert E. Kraut. “Effects of Head-Mounted and Scene-Oriented Video Systems on Remote Collaboration on Physical Tasks.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI ’03. Ft. Lauderdale, Florida, USA: Association for Computing Machinery, 2003, 513–520. ISBN: 1581136307. DOI: [10.1145/642611.642701](https://doi.org/10.1145/642611.642701). URL: <https://doi.org/10.1145/642611.642701>.

- [68] 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: *Hum.-Comput. Interact.* 19.3 (2004), 273–309. ISSN: 0737-0024. DOI: [10.1207/s15327051hci1903_3](https://doi.org/10.1207/s15327051hci1903_3). URL: https://doi.org/10.1207/s15327051hci1903_3.
- [69] Gerard Cesar Gabriel and Mary Lou Maher. "Coding and modelling communication in architectural collaborative design." In: *Automation in Construction* 11.2 (2002). ACADIA '99, pp. 199–211. ISSN: 0926-5805. DOI: [https://doi.org/10.1016/S0926-5805\(00\)00098-4](https://doi.org/10.1016/S0926-5805(00)00098-4). URL: <https://www.sciencedirect.com/science/article/pii/S0926580500000984>.
- [70] Lei Gao, Huidong Bai, Mark Billinghamurst, and Robert W. Lindeman. "User Behaviour Analysis of Mixed Reality Remote Collaboration with a Hybrid View Interface." In: *Proceedings of the 32nd Australian Conference on Human-Computer Interaction. OzCHI '20*. Sydney, NSW, Australia: Association for Computing Machinery, 2021, 629–638. ISBN: 9781450389754. DOI: [10.1145/3441000.3441038](https://doi.org/10.1145/3441000.3441038). URL: <https://doi.org/10.1145/3441000.3441038>.
- [71] Lei Gao, Huidong Bai, Weiping He, Mark Billinghamurst, and Robert W. Lindeman. "Real-Time Visual Representations for Mobile Mixed Reality Remote Collaboration." In: *SIGGRAPH Asia 2018 Virtual & Augmented Reality. SA '18*. Tokyo, Japan: Association for Computing Machinery, 2018. ISBN: 9781450360289. DOI: [10.1145/3275495.3275515](https://doi.org/10.1145/3275495.3275515). URL: <https://doi.org/10.1145/3275495.3275515>.
- [72] Lei Gao, Huidong Bai, Gun Lee, and Mark Billinghamurst. "An Oriented Point-Cloud View for MR Remote Collaboration." In: *SIGGRAPH ASIA 2016 Mobile Graphics and Interactive Applications. SA '16*. Macau: Association for Computing Machinery, 2016. ISBN: 9781450345514. DOI: [10.1145/2999508.2999531](https://doi.org/10.1145/2999508.2999531). URL: <https://doi.org/10.1145/2999508.2999531>.
- [73] Steffen Gauglitz, Cha Lee, Matthew Turk, and Tobias Höllerer. "Integrating the Physical Environment into Mobile Remote Collaboration." In: *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '12*. San Francisco, California, USA: Association for Computing Machinery, 2012, 241–250. ISBN: 9781450311052. DOI: [10.1145/2371574.2371610](https://doi.org/10.1145/2371574.2371610). URL: <https://doi.org/10.1145/2371574.2371610>.
- [74] Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. "In Touch with the Remote World: Remote Collaboration with Augmented Reality Drawings and Virtual Navigation." In: *Proceedings of the 20th ACM Symposium on*

- Virtual Reality Software and Technology*. VRST '14. Edinburgh, Scotland: Association for Computing Machinery, 2014, 197–205. ISBN: 9781450332538. DOI: [10.1145/2671015.2671016](https://doi.org/10.1145/2671015.2671016). URL: <https://doi.org/10.1145/2671015.2671016>.
- [75] William W. Gaver, Abigail Sellen, Christian Heath, and Paul Luff. “One is Not Enough: Multiple Views in a Media Space.” In: *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*. CHI '93. Amsterdam, The Netherlands: Association for Computing Machinery, 1993, 335–341. ISBN: 0897915755. DOI: [10.1145/169059.169268](https://doi.org/10.1145/169059.169268). URL: <https://doi.org/10.1145/169059.169268>.
- [76] Darren Gergle, Robert E. Kraut, and Susan R. Fussell. “Action as Language in a Shared Visual Space.” In: *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*. CSCW '04. Chicago, Illinois, USA: Association for Computing Machinery, 2004, 487–496. ISBN: 1581138105. DOI: [10.1145/1031607.1031687](https://doi.org/10.1145/1031607.1031687). URL: <https://doi.org/10.1145/1031607.1031687>.
- [77] Florian Geyer, Jochen Budzinski, and Harald Reiterer. “IdeaVis: A Hybrid Workspace and Interactive Visualization for Paper-Based Collaborative Sketching Sessions.” In: *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*. NordiCHI '12. Copenhagen, Denmark: Association for Computing Machinery, 2012, 331–340. ISBN: 9781450314824. DOI: [10.1145/2399016.2399069](https://doi.org/10.1145/2399016.2399069). URL: <https://doi.org/10.1145/2399016.2399069>.
- [78] Audrey Girouard, Orit Shaer, Erin T. Solovey, G. Michael Poor, and Robert J. K. Jacob. “The Reality of Reality-Based Interaction: Understanding the Impact of a Framework as a Research Tool.” In: *ACM Trans. Comput.-Hum. Interact.* 26.5 (2019). ISSN: 1073-0516. DOI: [10.1145/3319617](https://doi.org/10.1145/3319617). URL: <https://doi.org/10.1145/3319617>.
- [79] Daniele Giunchi, Stuart James, and Anthony Steed. “3D Sketching for Interactive Model Retrieval in Virtual Reality.” In: *Proceedings of the Joint Symposium on Computational Aesthetics and Sketch-Based Interfaces and Modeling and Non-Photorealistic Animation and Rendering*. Expressive '18. Victoria, British Columbia, Canada: Association for Computing Machinery, 2018. ISBN: 9781450358927. DOI: [10.1145/3229147.3229166](https://doi.org/10.1145/3229147.3229166). URL: <https://doi.org/10.1145/3229147.3229166>.
- [80] Leonardo Giusti, Kotval Xerxes, Amelia Schladow, Nicholas Wallen, Francis Zane, and Federico Casalegno. “Workspace Configurations: Setting the Stage for Remote Collaboration on Physical Tasks.” In: *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*.

- NordiCHI '12. Copenhagen, Denmark: Association for Computing Machinery, 2012, 351–360. ISBN: 9781450314824. DOI: [10.1145/2399016.2399071](https://doi.org/10.1145/2399016.2399071). URL: <https://doi.org/10.1145/2399016.2399071>.
- [81] Vinod Goel. *Sketches of Thought*. MIT Press, 1995.
- [82] Gabriela Goldschmidt. *Linkography: Unfolding the Design Process*. The MIT Press, Mar. 2014. ISBN: 9780262322157. DOI: [10.7551/mitpress/9455.001.0001](https://doi.org/10.7551/mitpress/9455.001.0001). URL: <https://doi.org/10.7551/mitpress/9455.001.0001>.
- [83] Jerônimo Gustavo Grandi, Henrique Galvan Debarba, and Anderson Maciel. “Characterizing Asymmetric Collaborative Interactions in Virtual and Augmented Realities.” In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2019, pp. 127–135. DOI: [10.1109/VR.2019.8798080](https://doi.org/10.1109/VR.2019.8798080).
- [84] Sydney Gregory. *The Design Method*. 1966.
- [85] Tovi Grossman and Ravin Balakrishnan. “Collaborative Interaction with Volumetric Displays.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '08. Florence, Italy: Association for Computing Machinery, 2008, 383–392. ISBN: 9781605580111. DOI: [10.1145/1357054.1357118](https://doi.org/10.1145/1357054.1357118). URL: <https://doi.org/10.1145/1357054.1357118>.
- [86] 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 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: Association for Computing Machinery, 2017, 4021–4033. ISBN: 9781450346559. DOI: [10.1145/3025453.3025683](https://doi.org/10.1145/3025453.3025683). URL: <https://doi.org/10.1145/3025453.3025683>.
- [87] Jake Guida and Misha Sra. “Augmented Reality World Editor.” In: *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*. VRST '20. Virtual Event, Canada: Association for Computing Machinery, 2020. ISBN: 9781450376198. DOI: [10.1145/3385956.3422125](https://doi.org/10.1145/3385956.3422125). URL: <https://doi.org/10.1145/3385956.3422125>.
- [88] Anhong Guo, Ilter Canberk, Hannah Murphy, Andrés Monroy-Hernández, and Rajan Vaish. “Blocks: Collaborative and Persistent Augmented Reality Experiences.” In: *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3.3 (2019). DOI: [10.1145/3351241](https://doi.org/10.1145/3351241). URL: <https://doi.org/10.1145/3351241>.
- [89] Carl Gutwin and Saul Greenberg. “A Descriptive Framework of Workspace Awareness for Real-Time Groupware.” In: *Computer Supported Cooperative Work (CSCW)* 11 (Sept. 2002), pp. 411–. DOI: [10.1023/A:1021271517844](https://doi.org/10.1023/A:1021271517844).

- [90] Martin Hachet, Fabrice Declé, Sebastian Knoedel, and Pascal Guitton. "Navidget for Easy 3D Camera Positioning from 2D Inputs." In: *Proceedings of the IEEE Symposium on 3D User Interfaces (3DUI)*. United States, 2008, pp. 83–88. URL: <https://hal.archives-ouvertes.fr/hal-00308251>.
- [91] Andreas Hackl and Helmut Hlavacs. "Diminishing Reality." In: *Entertainment Computing - ICEC 2018 - 17th IFIP TC 14 International Conference, Held at the 24th IFIP World Computer Congress, WCC 2018, Poznan, Poland, September 17-20, 2018, Proceedings*. 2018, pp. 28–39. URL: <http://eprints.cs.univie.ac.at/5891/>.
- [92] Joshua Hailpern, Erik Hinterbichler, Caryn Leppert, Damon Cook, and Brian P. Bailey. "TEAM STORM: Demonstrating an Interaction Model for Working with Multiple Ideas during Creative Group Work." In: *Proceedings of the 6th ACM SIGCHI Conference on Creativity & Cognition*. C&C '07. Washington, DC, USA: Association for Computing Machinery, 2007, 193–202. ISBN: 9781595937124. DOI: [10.1145/1254960.1254987](https://doi.org/10.1145/1254960.1254987). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/1254960.1254987>.
- [93] B. Hallgrímsson. *Prototyping and Modelmaking for Product Design*. Laurence King Publishing, 2012.
- [94] Steven R. Haynes, Amie L. Skattebo, Jonathan A. Singel, Mark A. Cohen, and Jodi L. Himelright. "Collaborative Architecture Design and Evaluation." In: *Proceedings of the 6th Conference on Designing Interactive Systems*. DIS '06. University Park, PA, USA: Association for Computing Machinery, 2006, 219–228. ISBN: 1595933670. DOI: [10.1145/1142405.1142440](https://doi.org/10.1145/1142405.1142440). URL: <https://doi.org/10.1145/1142405.1142440>.
- [95] Christian Heath and Paul Luff. "Disembodied conduct: Communication through video in a multi-media office environment." In: vol. 99-103. Jan. 1991, pp. 99–103. DOI: [10.1145/108844.108859](https://doi.org/10.1145/108844.108859).
- [96] Jon Hindmarsh, Mike Fraser, Christian Heath, Steve Benford, and Chris Greenhalgh. "Object-Focused Interaction in Collaborative Virtual Environments." In: *ACM Trans. Comput.-Hum. Interact.* 7.4 (2000), 477–509. ISSN: 1073-0516. DOI: [10.1145/365058.365088](https://doi.org/10.1145/365058.365088). URL: <https://doi.org/10.1145/365058.365088>.
- [97] 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: Association for Computing Machinery, 1992, 119–125. ISBN: 0897915135. DOI: [10.1145/142750.142769](https://doi.org/10.1145/142750.142769). URL: <https://doi.org/10.1145/142750.142769>.

- [98] Weidong Huang and Leila Alem. “HandsInAir: A Wearable System for Remote Collaboration on Physical Tasks.” In: *Proceedings of the 2013 Conference on Computer Supported Cooperative Work Companion*. CSCW ’13. San Antonio, Texas, USA: Association for Computing Machinery, 2013, 153–156. ISBN: 9781450313322. DOI: [10.1145/2441955.2441994](https://doi.org/10.1145/2441955.2441994). URL: <https://doi.org/10.1145/2441955.2441994>.
- [99] Ke Huo, Vinayak, and Karthik Ramani. “Window-Shaping: 3D Design Ideation by Creating on, Borrowing from, and Looking at the Physical World.” In: *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction*. TEI ’17. Yokohama, Japan: Association for Computing Machinery, 2017, 37–45. ISBN: 9781450346764. DOI: [10.1145/3024969.3024995](https://doi.org/10.1145/3024969.3024995). URL: <https://doi.org/10.1145/3024969.3024995>.
- [100] Ke Huo, Tianyi Wang, Luis Paredes, Ana M. Villanueva, Yuanzhi Cao, and Karthik Ramani. “SynchronizAR: Instant Synchronization for Spontaneous and Spatial Collaborations in Augmented Reality.” In: *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. UIST ’18. Berlin, Germany: Association for Computing Machinery, 2018, 19–30. ISBN: 9781450359481. DOI: [10.1145/3242587.3242595](https://doi.org/10.1145/3242587.3242595). URL: <https://doi.org/10.1145/3242587.3242595>.
- [101] Ananya Ipsita, Hao Li, Runlin Duan, Yuanzhi Cao, Subramanian Chidambaram, Min Liu, and Karthik Ramani. “VRFromX: From Scanned Reality to Interactive Virtual Experience with Human-in-the-Loop.” In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI EA ’21. Yokohama, Japan: Association for Computing Machinery, 2021. ISBN: 9781450380959. DOI: [10.1145/3411763.3451747](https://doi-org.ins2i.bib.cnrs.fr/10.1145/3411763.3451747). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3411763.3451747>.
- [102] Ellen A. Isaacs and John C. Tang. “What Video Can and Can’t Do for Collaboration: A Case Study.” In: *Proceedings of the First ACM International Conference on Multimedia*. MULTIMEDIA ’93. Anaheim, California, USA: Association for Computing Machinery, 1993, 199–206. ISBN: 0897915968. DOI: [10.1145/166266.166289](https://doi.org/10.1145/166266.166289). URL: <https://doi.org/10.1145/166266.166289>.
- [103] Hiroshi Ishii, Minoru Kobayashi, and Kazuho Arita. “Iterative Design of Seamless Collaboration Media.” In: *Commun. ACM* 37.8 (1994), 83–97. ISSN: 0001-0782. DOI: [10.1145/179606.179687](https://doi.org/10.1145/179606.179687). URL: <https://doi.org/10.1145/179606.179687>.
- [104] Shahram Izadi et al. “KinectFusion: Real-Time 3D Reconstruction and Interaction Using a Moving Depth Camera.” In: *Proceedings of the 24th Annual ACM Symposium on User Interface*

- Software and Technology*. UIST '11. Santa Barbara, California, USA: Association for Computing Machinery, 2011, 559–568. ISBN: 9781450307161. DOI: [10.1145/2047196.2047270](https://doi.org/10.1145/2047196.2047270). URL: <https://doi.org/10.1145/2047196.2047270>.
- [105] 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 33rd Annual ACM Symposium on User Interface Software and Technology*. UIST '20. Virtual Event, USA: Association for Computing Machinery, 2020, 459–471. ISBN: 9781450375146. DOI: [10.1145/3379337.3415843](https://doi.org/10.1145/3379337.3415843). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3379337.3415843>.
- [106] Allison Jing, Kieran William May, Mahnoor Naeem, Gun Lee, and Mark Billingham. “EyemR-Vis: Using Bi-Directional Gaze Behavioural Cues to Improve Mixed Reality Remote Collaboration.” In: *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI EA '21. Yokohama, Japan: Association for Computing Machinery, 2021. ISBN: 9781450380959. DOI: [10.1145/3411763.3451844](https://doi.org/10.1145/3411763.3451844). URL: <https://doi.org/10.1145/3411763.3451844>.
- [107] Allison Jing, Kieran May, Brandon Matthews, Gun Lee, and Mark Billingham. “The Impact of Sharing Gaze Behaviours in Collaborative Mixed Reality.” In: *Proc. ACM Hum.-Comput. Interact.* 6.CSCW2 (2022). DOI: [10.1145/3555564](https://doi.org/10.1145/3555564). URL: <https://doi.org/10.1145/3555564>.
- [108] Peter Johnson-Lenz and Trudy Johnson-Lenz. “Post-mechanistic groupware primitives: rhythms, boundaries and containers.” In: *International Journal of Man-Machine Studies* 34.3 (1991). Computer-supported Cooperative Work and Groupware. Part 2, pp. 395–417. ISSN: 0020-7373. DOI: [https://doi.org/10.1016/0020-7373\(91\)90027-5](https://doi.org/10.1016/0020-7373(91)90027-5). URL: <https://www.sciencedirect.com/science/article/pii/0020737391900275>.
- [109] Brennan Jones, Anna Witcraft, Scott Bateman, Carman Neustaedter, and Anthony Tang. “Mechanics of Camera Work in Mobile Video Collaboration.” In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. Seoul, Republic of Korea: Association for Computing Machinery, 2015, 957–966. ISBN: 9781450331456. DOI: [10.1145/2702123.2702345](https://doi.org/10.1145/2702123.2702345). URL: <https://doi.org/10.1145/2702123.2702345>.
- [110] Brennan Jones, Yaying Zhang, Priscilla N. Y. Wong, and Sean Rintel. “Belonging There: VROOM-Ing into the Uncanny Valley of XR Telepresence.” In: *Proc. ACM Hum.-Comput. Interact.* 5.CSCW1 (2021). DOI: [10.1145/3449133](https://doi.org/10.1145/3449133). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3449133>.

- [111] Nicolas Kahrl, Michael Prilla, and Oliver Blunk. "Show Me Your Living Room: Investigating the Role of Representing User Environments in AR Remote Consultations." In: *Proceedings of the Conference on Mensch Und Computer*. MuC '20. Magdeburg, Germany: Association for Computing Machinery, 2020, 267–277. ISBN: 9781450375405. DOI: [10.1145/3404983.3405520](https://doi.org/10.1145/3404983.3405520). URL: <https://doi.org/10.1145/3404983.3405520>.
- [112] Denis Kalkofen, Eduardo Veas, Stefanie Zollmann, Markus Steinberger, and Dieter Schmalstieg. "Adaptive Ghosted Views for Augmented Reality." In: Oct. 2013. DOI: [10.1109/ISMAR.2013.6671758](https://doi.org/10.1109/ISMAR.2013.6671758).
- [113] T. Kanade, P. Rander, and P.J. Narayanan. "Virtualized reality: constructing virtual worlds from real scenes." In: *IEEE Multi-Media* 4.1 (1997), pp. 34–47. ISSN: 1941-0166. DOI: [10.1109/93.580394](https://doi.org/10.1109/93.580394).
- [114] Zhizhong Kang, Juntao Yang, Zhou Yang, and Sai Cheng. "A Review of Techniques for 3D Reconstruction of Indoor Environments." In: *ISPRS International Journal of Geo-Information* 9.5 (2020). ISSN: 2220-9964. DOI: [10.3390/ijgi9050330](https://doi.org/10.3390/ijgi9050330). URL: <https://www.mdpi.com/2220-9964/9/5/330>.
- [115] Shunichi Kasahara, Valentin Heun, Austin S. Lee, and Hiroshi Ishii. "Second Surface: Multi-User Spatial Collaboration System Based on Augmented Reality." In: *SIGGRAPH Asia 2012 Emerging Technologies*. SA '12. Singapore, Singapore: Association for Computing Machinery, 2012, 1–4. ISBN: 9781450319126. DOI: [10.1145/2407707.2407727](https://doi.org/10.1145/2407707.2407727). URL: <https://doi.org/10.1145/2407707.2407727>.
- [116] Taiki Kato, Naoya Isoyama, Norihiko Kawai, Hideaki Uchiyama, Nobuchika Sakata, and Kiyoshi Kiyokawa. "Online Adaptive Integration of Observation and Inpainting for Diminished Reality with Online Surface Reconstruction." In: *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 2022, pp. 308–314. DOI: [10.1109/ISMAR-Adjunct57072.2022.00069](https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00069).
- [117] Randy H Katz, Rajiv Bhateja, Ellis E-li Chang, David Gedye, and Vony Trijanto. "Design Version Management." In: *IEEE Design & Test of Computers* 4.1 (1987), pp. 12–22. DOI: [10.1109/MDT.1987.295109](https://doi.org/10.1109/MDT.1987.295109).
- [118] Hannes Kaufmann and Dieter Schmalstieg. "Mathematics and Geometry Education with Collaborative Augmented Reality." In: *ACM SIGGRAPH 2002 Conference Abstracts and Applications*. SIGGRAPH '02. San Antonio, Texas: Association for Computing Machinery, 2002, 37–41. ISBN: 1581135254. DOI: [10.1145/1242073.1242086](https://doi.org/10.1145/1242073.1242086). URL: <https://doi.org/10.1145/1242073.1242086>.

- [119] Matthew Kay, Gregory L. Nelson, and Eric B. Hekler. "Researcher-Centered Design of Statistics: Why Bayesian Statistics Better Fit the Culture and Incentives of HCI." In: CHI '16. San Jose, California, USA: Association for Computing Machinery, 2016, 4521–4532. ISBN: 9781450333627. DOI: [10.1145/2858036.2858465](https://doi.org/10.1145/2858036.2858465). URL: <https://doi.org/10.1145/2858036.2858465>.
- [120] 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 30th Annual ACM Symposium on User Interface Software and Technology*. UIST '17. Québec City, QC, Canada: Association for Computing Machinery, 2017, 401–414. ISBN: 9781450349819. DOI: [10.1145/3126594.3126662](https://doi.org/10.1145/3126594.3126662). URL: <https://doi.org/10.1145/3126594.3126662>.
- [121] Cedric Kervegant, Julien Castet, Juliette Vauchez, and Charles Bailly. "Distant Assist Cursor (DAC): Designing an Augmented Reality System to Facilitate Remote Collaboration for Novice Users." In: *Companion Proceedings of the 2021 Conference on Interactive Surfaces and Spaces*. ISS '21. Lodz, Poland: Association for Computing Machinery, 2021, 8–11. ISBN: 9781450383400. DOI: [10.1145/3447932.3490520](https://doi.org/10.1145/3447932.3490520). URL: <https://doi.org/10.1145/3447932.3490520>.
- [122] Seungwon Kim, Gun A. Lee, and Nobuchika Sakata. "Comparing pointing and drawing for remote collaboration." In: *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2013, pp. 1–6. DOI: [10.1109/ISMAR.2013.6671833](https://doi.org/10.1109/ISMAR.2013.6671833).
- [123] Seungwon Kim, Gun Lee, Nobuchika Sakata, and Mark Billinghurst. "Improving co-presence with augmented visual communication cues for sharing experience through video conference." In: *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 2014, pp. 83–92. DOI: [10.1109/ISMAR.2014.6948412](https://doi.org/10.1109/ISMAR.2014.6948412).
- [124] David Kirk, Andy Crabtree, and Tom Rodden. "Ways of the hands." In: *ECSCW 2005*. Springer, 2005, pp. 1–21.
- [125] David Kirk and Danae Stanton Fraser. "Comparing Remote Gesture Technologies for Supporting Collaborative Physical Tasks." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '06. Montréal, Québec, Canada: Association for Computing Machinery, 2006, 1191–1200. ISBN: 1595933727. DOI: [10.1145/1124772.1124951](https://doi.org/10.1145/1124772.1124951). URL: <https://doi.org/10.1145/1124772.1124951>.
- [126] K. Kiyokawa, M. Billinghurst, S. E. Hayes, A. Gupta, Y. Sannohe, and H. Kato. "Communication Behaviors of Co-Located Users in Collaborative AR Interfaces." In: *Proceedings of the 1st Inter-*

- national Symposium on Mixed and Augmented Reality*. ISMAR '02. USA: IEEE Computer Society, 2002, p. 139. ISBN: 0769517811.
- [127] K. Kiyokawa, H. Takemura, and N. Yokoya. "A collaboration support technique by integrating a shared virtual reality and a shared augmented reality." In: *IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No.99CH37028)*. Vol. 6. 1999, 48–53 vol.6. DOI: [10.1109/ICSMC.1999.816444](https://doi.org/10.1109/ICSMC.1999.816444).
- [128] Kiyoshi Kiyokawa, Haruo Takemura, and Naokazu Yokoya. "SeamlessDesign for 3D Object Creation." In: *IEEE MultiMedia* 7.1 (2000), 22–33. ISSN: 1070-986X. DOI: [10.1109/93.839308](https://doi.org/10.1109/93.839308). URL: <https://doi.org/10.1109/93.839308>.
- [129] Ryohei Komiyama, Takashi Miyaki, and Jun Rekimoto. "JackIn Space: Designing a Seamless Transition between First and Third Person View for Effective Telepresence Collaborations." In: *Proceedings of the 8th Augmented Human International Conference*. AH '17. Silicon Valley, California, USA: Association for Computing Machinery, 2017. ISBN: 9781450348355. DOI: [10.1145/3041164.3041183](https://doi.org/10.1145/3041164.3041183). URL: <https://doi.org/10.1145/3041164.3041183>.
- [130] Dimosthenis Kontogiorgos, Elena Sibirtseva, Andre Pereira, Gabriel Skantze, and Joakim Gustafson. "Multimodal Reference Resolution In Collaborative Assembly Tasks." In: *Proceedings of the 4th International Workshop on Multimodal Analyses Enabling Artificial Agents in Human-Machine Interaction*. MA3HMI'18. Boulder, CO, USA: Association for Computing Machinery, 2018, 38–42. ISBN: 9781450360760. DOI: [10.1145/3279972.3279976](https://doi.org/10.1145/3279972.3279976). URL: <https://doi.org/10.1145/3279972.3279976>.
- [131] Anna K. Kuhlen and Susan E. Brennan. "Language in dialogue: when confederates might be hazardous to your data." In: *Psychonomic Bulletin & Review* 20.1 (2013), pp. 54–72. DOI: [10.3758/s13423-012-0341-8](https://doi.org/10.3758/s13423-012-0341-8). URL: <https://doi.org/10.3758/s13423-012-0341-8>.
- [132] André Kunert, Alexander Kulik, Stephan Beck, and Bernd Froehlich. "Photoportals: Shared References in Space and Time." In: *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*. CSCW '14. Baltimore, Maryland, USA: Association for Computing Machinery, 2014, 1388–1399. ISBN: 9781450325400. DOI: [10.1145/2531602.2531727](https://doi.org/10.1145/2531602.2531727). URL: <https://doi.org/10.1145/2531602.2531727>.
- [133] Hideaki Kuzuoka. "Spatial Workspace Collaboration: A Shared-View Video Support System for Remote Collaboration Capability." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '92. Monterey, California, USA: Association for Computing Machinery, 1992, 533–540.

- ISBN: 0897915135. DOI: [10.1145/142750.142980](https://doi.org/10.1145/142750.142980). URL: <https://doi.org/10.1145/142750.142980>.
- [134] Hideaki Kuzuoka, Toshio Kosuge, and Masatomo Tanaka. "GestureCam: A Video Communication System for Sympathetic Remote Collaboration." In: *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*. CSCW '94. Chapel Hill, North Carolina, USA: Association for Computing Machinery, 1994, 35–43. ISBN: 0897916891. DOI: [10.1145/192844.192866](https://doi.org/10.1145/192844.192866). URL: <https://doi.org/10.1145/192844.192866>.
- [135] Kin Chung Kwan and Hongbo Fu. "Mobi3DSketch: 3D Sketching in Mobile AR." In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019, 1–11. ISBN: 9781450359702. DOI: [10.1145/3290605.3300406](https://doi.org/10.1145/3290605.3300406). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3290605.3300406>.
- [136] James A. Landay and Brad A. Myers. "Interactive Sketching for the Early Stages of User Interface Design." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '95. Denver, Colorado, USA: ACM Press/Addison-Wesley Publishing Co., 1995, 43–50. ISBN: 0201847051. DOI: [10.1145/223904.223910](https://doi.org/10.1145/223904.223910). URL: <https://doi.org/10.1145/223904.223910>.
- [137] Joel Lanir, Ran Stone, Benjamin Cohen, and Pavel Gurevich. "Ownership and Control of Point of View in Remote Assistance." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2013, 2243–2252. ISBN: 9781450318990. URL: <https://doi.org/10.1145/2470654.2481309>.
- [138] Joseph Laviola, Loring Holden, Andrew Forsberg, Dom Bhuphaibool, and Robert Zeleznik. "Collaborative Conceptual Modeling Using The Sketch Framework." In: (June 2000).
- [139] Jeremy Laviolle and Martin Hachet. "Demo: Spatial augmented reality for physical drawing." In: Oct. 2012, pp. 9–10. DOI: [10.1145/2380296.2380302](https://doi.org/10.1145/2380296.2380302).
- [140] Bryan Lawson. "How Designers Think – The Design Process Demystified." In: *University Press, Cambridge* (Jan. 2006).
- [141] Bryan Lawson. *What Designers Know*. Oct. 2013.
- [142] Khanh-Duy Le, Paweł W. Woźniak, Ali Alavi, Morten Fjeld, and Andreas Kunz. "DigiMetaplan: Supporting Facilitated Brainstorming for Distributed Business Teams." In: *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*. MUM '19. Pisa, Italy: Association for Computing Machinery, 2019. ISBN: 9781450376242. DOI: [10.1145/3365610.3365637](https://doi.org/10.1145/3365610.3365637).

- URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3365610.3365637>.
- [143] Gun A. Lee, Theophilus Teo, Seungwon Kim, and Mark Billinghurst. "Sharedsphere: MR Collaboration through Shared Live Panorama." In: *SIGGRAPH Asia 2017 Emerging Technologies*. SA '17. Bangkok, Thailand: Association for Computing Machinery, 2017. ISBN: 9781450354042. DOI: [10.1145/3132818.3132827](https://doi.org/10.1145/3132818.3132827). URL: <https://doi.org/10.1145/3132818.3132827>.
- [144] Seunghyun Lee and Ellen Yi-Luen Do. "The Effects of Computing Technology in Creative Design Tasks: A Case Study of Design Collaboration." In: *Proceedings of the Seventh ACM Conference on Creativity and Cognition*. C&C '09. Berkeley, California, USA: Association for Computing Machinery, 2009, 387–388. ISBN: 9781605588650. DOI: [10.1145/1640233.1640314](https://doi.org/10.1145/1640233.1640314). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/1640233.1640314>.
- [145] Yuan Li, David Hicks, Wallace S. Lages, Sang Won Lee, Akshay Sharma, and Doug A. Bowman. "ARCritique: Supporting Remote Design Critique of Physical Artifacts through Collaborative Augmented Reality." In: *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 2021, pp. 585–586. DOI: [10.1109/VRW52623.2021.00175](https://doi.org/10.1109/VRW52623.2021.00175).
- [146] Torrin M. Liddell and John K. Kruschke. "Analyzing ordinal data with metric models: What could possibly go wrong?" In: *Journal of Experimental Social Psychology* 79 (2018), pp. 328–348. ISSN: 0022-1031. DOI: <https://doi.org/10.1016/j.jesp.2018.08.009>. URL: <https://www.sciencedirect.com/science/article/pii/S0022103117307746>.
- [147] Klemen Lilija, Henning Pohl, Sebastian Boring, and Kasper Hornbæk. "Augmented Reality Views for Occluded Interaction." In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019, 1–12. ISBN: 9781450359702. DOI: [10.1145/3290605.3300676](https://doi.org/10.1145/3290605.3300676). URL: <https://doi.org/10.1145/3290605.3300676>.
- [148] Yishuo Liu, Yichuan Zhang, Shiliang Zuo, and Wai-Tat Fu. "BoatAR: A Multi-User Augmented-Reality Platform for Boat." In: *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. VRST '18. Tokyo, Japan: Association for Computing Machinery, 2018. ISBN: 9781450360869. DOI: [10.1145/3281505.3283392](https://doi.org/10.1145/3281505.3283392). URL: <https://doi.org/10.1145/3281505.3283392>.

- [149] Dastyni Loksa, Nicolas Mangano, Thomas D. LaToza, and André van der Hoek. “Enabling a Classroom Design Studio with a Collaborative Sketch Design Tool.” In: *Proceedings of the 2013 International Conference on Software Engineering*. ICSE '13. San Francisco, CA, USA: IEEE Press, 2013, 1073–1082. ISBN: 9781467330763.
- [150] Julian Looser, Mark Billingham, and Andy Cockburn. “Through the Looking Glass: The Use of Lenses as an Interface Tool for Augmented Reality Interfaces.” In: *Proceedings of the 2nd International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia*. GRAPHITE '04. Singapore: Association for Computing Machinery, 2004, 204–211. ISBN: 1581138830. DOI: [10.1145/988834.988870](https://doi.org/10.1145/988834.988870). URL: <https://doi.org/10.1145/988834.988870>.
- [151] Julian Looser, Mark Billingham, Raphaël Grasset, and Andy Cockburn. “An Evaluation of Virtual Lenses for Object Selection in Augmented Reality.” In: *Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia*. GRAPHITE '07. Perth, Australia: Association for Computing Machinery, 2007, 203–210. ISBN: 9781595939128. DOI: [10.1145/1321261.1321297](https://doi.org/10.1145/1321261.1321297). URL: <https://doi.org/10.1145/1321261.1321297>.
- [152] Julian Looser, Raphael Grasset, and Mark Billingham. “A 3D Flexible and Tangible Magic Lens in Augmented Reality.” In: Dec. 2007, pp. 51–54. ISBN: 978-1-4244-1749-0. DOI: [10.1109/ISMAR.2007.4538825](https://doi.org/10.1109/ISMAR.2007.4538825).
- [153] Xuan Luo, Jia-Bin Huang, Richard Szeliski, Kevin Matzen, and Johannes Kopf. “Consistent Video Depth Estimation.” In: *ACM Trans. Graph.* 39.4 (2020). ISSN: 0730-0301. DOI: [10.1145/3386569.3392377](https://doi.org/10.1145/3386569.3392377). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3386569.3392377>.
- [154] Zhuoyue Lyu, Jackie (Junrui) Yang, Monica S. Lam, and James A. Landay. “HomeView: Automatically Building Smart Home Digital Twins With Augmented Reality Headsets.” In: *The Adjunct Publication of the 35th Annual ACM Symposium on User Interface Software and Technology*. UIST '22 Adjunct. Bend, OR, USA: Association for Computing Machinery, 2022. ISBN: 97814503932-18. DOI: [10.1145/3526114.3558709](https://doi.org/10.1145/3526114.3558709). URL: <https://doi.org/10.1145/3526114.3558709>.
- [155] Eva Babette Mackamul and Augusto Esteves. “A Look at the Effects of Handheld and Projected Augmented-Reality on a Collaborative Task.” In: *Proceedings of the 2018 ACM Symposium on Spatial User Interaction*. SUI '18. Berlin, Germany: Association for Computing Machinery, 2018, 74–78. ISBN: 9781450357081.

- DOI: [10.1145/3267782.3267793](https://doi.org/10.1145/3267782.3267793). URL: <https://doi.org/10.1145/3267782.3267793>.
- [156] E. Marchand, F. Spindler, and F. Chaumette. “ViSP for visual servoing: a generic software platform with a wide class of robot control skills.” In: *IEEE Robotics and Automation Magazine* 12.4 (2005), pp. 40–52.
- [157] Swann Martinez and Chu-Yin Chen. “A Framework Enabling Real-time Multi-user Collaborative Workflow in 3D Digital Content Creation Software.” In: Jan. 2021, pp. 91–100. DOI: [10.24132/CSRN.2021.3002.10](https://doi.org/10.24132/CSRN.2021.3002.10).
- [158] Paul W. Mattessich and Barbara R. Monsey. “Collaboration: What Makes It Work- A Review of Research Literature on Factors Influencing Successful Collaboration.” In: 1992.
- [159] Sven Mayer, Jens Reinhardt, Robin Schweigert, Brighten Jelke, Valentin Schwind, Katrin Wolf, and Niels Henze. “Improving Humans’ Ability to Interpret Deictic Gestures in Virtual Reality.” In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI ’20. Honolulu, HI, USA: Association for Computing Machinery, 2020, 1–14. ISBN: 9781450367080. DOI: [10.1145/3313831.3376340](https://doi.org/10.1145/3313831.3376340). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3313831.3376340>.
- [160] M.J. McGuffin, L. Tancau, and R. Balakrishnan. “Using deformations for browsing volumetric data.” In: *IEEE Visualization, 2003. VIS 2003*. 2003, pp. 401–408. DOI: [10.1109/VISUAL.2003.1250400](https://doi.org/10.1109/VISUAL.2003.1250400).
- [161] Liang Men and Nick Bryan-Kinns. “LeMo: Exploring Virtual Space for Collaborative Creativity.” In: *Proceedings of the 2019 on Creativity and Cognition*. C&C ’19. San Diego, CA, USA: Association for Computing Machinery, 2019, 71–82. ISBN: 9781450359177. DOI: [10.1145/3325480.3325495](https://doi.org/10.1145/3325480.3325495). URL: <https://doi.org/10.1145/3325480.3325495>.
- [162] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. “Augmented reality: a class of displays on the reality-virtuality continuum.” In: *Telem manipulator and Telepresence Technologies*. Ed. by Hari Das. Vol. 2351. International Society for Optics and Photonics. SPIE, 1995, pp. 282–292. URL: <https://doi.org/10.1117/12.197321>.
- [163] 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 2020 CHI Conference on Human Factors in Computing Systems*. CHI ’20. Honolulu, HI, USA: Association for Computing Machinery, 2020, 1–12. ISBN: 9781450367080. DOI: [10.1145/3313831.3376289](https://doi.org/10.1145/3313831.3376289). URL: <https://doi.org/10.1145/3313831.3376289>.

- [164] Jens Müller, Roman Rädle, and Harald Reiterer. "Virtual Objects as Spatial Cues in Collaborative Mixed Reality Environments: How They Shape Communication Behavior and User Task Load." In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. CHI '16. San Jose, California, USA: Association for Computing Machinery, 2016, 1245–1249. ISBN: 9781450333627. DOI: [10.1145/2858036.2858043](https://doi.org/10.1145/2858036.2858043). URL: <https://doi.org/10.1145/2858036.2858043>.
- [165] Jens Müller, Roman Rädle, and Harald Reiterer. "Remote Collaboration With Mixed Reality Displays: How Shared Virtual Landmarks Facilitate Spatial Referencing." In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: Association for Computing Machinery, 2017, 6481–6486. ISBN: 9781450346559. DOI: [10.1145/3025453.3025717](https://doi.org/10.1145/3025453.3025717). URL: <https://doi.org/10.1145/3025453.3025717>.
- [166] David T. Nguyen and John Canny. "Multiview: Improving Trust in Group Video Conferencing through Spatial Faithfulness." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '07. San Jose, California, USA: Association for Computing Machinery, 2007, 1465–1474. ISBN: 9781595935939. DOI: [10.1145/1240624.1240846](https://doi.org/10.1145/1240624.1240846). URL: <https://doi.org/10.1145/1240624.1240846>.
- [167] Hiroaki Nishino, Kouichi Utsumiya, Kazuyoshi Korida, Atsunori Sakamoto, and Kazuyuki Yoshida. "A Method for Sharing Interactive Deformations in Collaborative 3D Modeling." In: *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. VRST '99. London, United Kingdom: Association for Computing Machinery, 1999, 116–123. ISBN: 1581131410. DOI: [10.1145/323663.323682](https://doi.org/10.1145/323663.323682). URL: <https://doi.org/10.1145/323663.323682>.
- [168] Ohan Oda, Carmine Elvezio, Mengu Sukan, Steven Feiner, and Barbara Tversky. "Virtual Replicas for Remote Assistance in Virtual and Augmented Reality." In: *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. UIST '15. Charlotte, NC, USA: Association for Computing Machinery, 2015, 405–415. ISBN: 9781450337793. DOI: [10.1145/2807442.2807497](https://doi.org/10.1145/2807442.2807497). URL: <https://doi.org/10.1145/2807442.2807497>.
- [169] Yujiro Okuya, Olivier Gladin, Nicolas Ladeveze, Cédric Fleury, and Patrick Bourdot. "Investigating Collaborative Exploration of Design Alternatives on a Wall-Sized Display." In: Apr. 2020, pp. 1–12. DOI: [10.1145/3313831.3376736](https://doi.org/10.1145/3313831.3376736).

- [170] Niklas Osmers and Michael Prilla. "Getting out of Out of Sight: Evaluation of AR Mechanisms for Awareness and Orientation Support in Occluded Multi-Room Settings." In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Honolulu, HI, USA: Association for Computing Machinery, 2020, 1–11. ISBN: 9781450367080. DOI: [10.1145/3313831.3376742](https://doi.org/10.1145/3313831.3376742). URL: <https://doi.org/10.1145/3313831.3376742>.
- [171] Niklas Osmers, Michael Prilla, Oliver Blunk, Gordon George Brown, Marc Janßen, and Nicolas Kahrl. "The Role of Social Presence for Cooperation in Augmented Reality on Head Mounted Devices: A Literature Review." In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. CHI '21. Yokohama, Japan: Association for Computing Machinery, 2021. ISBN: 9781450380966. DOI: [10.1145/3411764.3445633](https://doi.org/10.1145/3411764.3445633). URL: <https://doi.org/10.1145/3411764.3445633>.
- [172] Jiazhi Ou, Lui Min Oh, Jie Yang, and Susan R. Fussell. "Effects of Task Properties, Partner Actions, and Message Content on Eye Gaze Patterns in a Collaborative Task." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '05. Portland, Oregon, USA: Association for Computing Machinery, 2005, 231–240. ISBN: 1581139985. DOI: [10.1145/1054972.1055005](https://doi.org/10.1145/1054972.1055005). URL: <https://doi.org/10.1145/1054972.1055005>.
- [173] Kyoung S. Park, Abhinav Kapoor, and Jason Leigh. "Lessons Learned from Employing Multiple Perspectives in a Collaborative Virtual Environment for Visualizing Scientific Data." In: *Proceedings of the Third International Conference on Collaborative Virtual Environments*. CVE '00. San Francisco, California, USA: Association for Computing Machinery, 2000, 73–82. ISBN: 1581133030. DOI: [10.1145/351006.351015](https://doi.org/10.1145/351006.351015). URL: <https://doi.org/10.1145/351006.351015>.
- [174] Sihwa Park. "ARLooper: Collaborative Audiovisual Experience with Mobile Devices in a Shared Augmented Reality Space." In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI EA '20. Honolulu, HI, USA: Association for Computing Machinery, 2020, 1–4. ISBN: 9781450368193. DOI: [10.1145/3334480.3383172](https://doi.org/10.1145/3334480.3383172). URL: <https://doi.org/10.1145/3334480.3383172>.
- [175] Cyprien Pindat, Emmanuel Pietriga, Oliver Chapuis, and Claude Puech. "Drilling into Complex 3D Models with Gimlenses." In: *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*. VRST '13. Singapore: Association for Computing Machinery, 2013, 223–230. ISBN: 9781450323796. DOI: [10.1145/2503713.2503714](https://doi.org/10.1145/2503713.2503714). URL: <https://doi.org/10.1145/2503713.2503714>.

- [176] David Pinelle, Carl Gutwin, and Saul Greenberg. "Task Analysis for Groupware Usability Evaluation: Modeling Shared-Workspace Tasks with the Mechanics of Collaboration." In: *ACM Trans. Comput.-Hum. Interact.* 10.4 (2003), 281–311. ISSN: 1073-0516. DOI: [10.1145/966930.966932](https://doi.org/10.1145/966930.966932). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/966930.966932>.
- [177] Vincent Adelizzi Pisani, Ocean Hurd, Nico Hawthorne, and Sri Kurniawan. "Navigation by Walking In Hyperbolic Space Using Virtual Reality." In: *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*. CHI PLAY '19 Extended Abstracts. Barcelona, Spain: Association for Computing Machinery, 2019, 611–618. ISBN: 9781450368711. DOI: [10.1145/3341215.3356287](https://doi.org/10.1145/3341215.3356287). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3341215.3356287>.
- [178] Thammathip Piumsomboon, Gun A. Lee, and Mark Billingham. "Snow Dome: A Multi-Scale Interaction in Mixed Reality Remote Collaboration." In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI EA '18. Montreal QC, Canada: Association for Computing Machinery, 2018, 1–4. ISBN: 9781450356213. DOI: [10.1145/3170427.3186495](https://doi.org/10.1145/3170427.3186495). URL: <https://doi.org/10.1145/3170427.3186495>.
- [179] Thammathip Piumsomboon, Gun A. Lee, Jonathon D. Hart, Barrett Ens, Robert W. Lindeman, Bruce H. Thomas, and Mark Billingham. "Mini-Me: An Adaptive Avatar for Mixed Reality Remote Collaboration." In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI '18. Montreal QC, Canada: Association for Computing Machinery, 2018, 1–13. ISBN: 9781450356206. DOI: [10.1145/3173574.3173620](https://doi.org/10.1145/3173574.3173620). URL: <https://doi.org/10.1145/3173574.3173620>.
- [180] Thammathip Piumsomboon, Gun A. Lee, Andrew Irlitti, Barrett Ens, Bruce H. Thomas, and Mark Billingham. "On the Shoulder of the Giant: A Multi-Scale Mixed Reality Collaboration with 360 Video Sharing and Tangible Interaction." In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019, 1–17. ISBN: 9781450359702. DOI: [10.1145/3290605.3300458](https://doi.org/10.1145/3290605.3300458). URL: <https://doi.org/10.1145/3290605.3300458>.
- [181] Thammathip Piumsomboon, Youngho Lee, Gun Lee, and Mark Billingham. "CoVAR: A Collaborative Virtual and Augmented Reality System for Remote Collaboration." In: *SIGGRAPH Asia 2017 Emerging Technologies*. SA '17. Bangkok, Thailand: Association for Computing Machinery, 2017. ISBN: 9781450354042. DOI: [10.1145/3132818.3132822](https://doi.org/10.1145/3132818.3132822). URL: <https://doi.org/10.1145/3132818.3132822>.

- [182] Cecil Piya, Vinayak, Senthil Chandrasegaran, Niklas Elmqvist, and Karthik Ramani. "Co-3Deator: A Team-First Collaborative 3D Design Ideation Tool." In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: Association for Computing Machinery, 2017, 6581–6592. ISBN: 9781450346559. DOI: [10.1145/3025453.3025825](https://doi.org/10.1145/3025453.3025825). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3025453.3025825>.
- [183] Pragathi Praveena, Luis Molina, Yeping Wang, Emmanuel Senft, Bilge Mutlu, and Michael Gleicher. "Understanding Control Frames in Multi-Camera Robot Telemanipulation." In: *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '22. Sapporo, Hokkaido, Japan: IEEE Press, 2022, 432–440.
- [184] Irene Rae, Bilge Mutlu, and Leila Takayama. "Bodies in Motion: Mobility, Presence, and Task Awareness in Telepresence." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada: Association for Computing Machinery, 2014, 2153–2162. ISBN: 9781450324731. DOI: [10.1145/2556288.2557047](https://doi.org/10.1145/2556288.2557047). URL: <https://doi.org/10.1145/2556288.2557047>.
- [185] Claude Raffestin and Samuel A Butler. "Space, Territory, and Territoriality." In: *Environment and Planning D: Society and Space* 30.1 (2012), pp. 121–141. DOI: [10.1068/d21311](https://doi.org/10.1068/d21311). eprint: <https://doi.org/10.1068/d21311>. URL: <https://doi.org/10.1068/d21311>.
- [186] Abhishek Ranjan, Jeremy P. Birnholtz, and Ravin Balakrishnan. "Dynamic Shared Visual Spaces: Experimenting with Automatic Camera Control in a Remote Repair Task." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2007, 1177–1186. ISBN: 9781595935939. URL: <https://doi.org/10.1145/1240624.1240802>.
- [187] Troels A. Rasmussen and Weidong Huang. "SceneCam: Using AR to Improve Multi-Camera Remote Collaboration." In: *SIGGRAPH Asia 2019 XR*. SA '19. Brisbane, QLD, Australia: Association for Computing Machinery, 2019, 36–37. ISBN: 9781450369473. DOI: [10.1145/3355355.3361892](https://doi.org/10.1145/3355355.3361892). URL: <https://doi.org/10.1145/3355355.3361892>.
- [188] Holger T. Regenbrecht and Michael T. Wagner. "Interaction in a Collaborative Augmented Reality Environment." In: *CHI '02 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '02. Minneapolis, Minnesota, USA: Association for Computing Machinery, 2002, 504–505. ISBN: 1581134541. DOI:

- 10.1145/506443.506451. URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/506443.506451>.
- [189] Patrick Reipschläger and Raimund Dachsel. “DesignAR: Immersive 3D-Modeling Combining Augmented Reality with Interactive Displays.” In: *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*. ISS ’19. Daejeon, Republic of Korea: Association for Computing Machinery, 2019, 29–41. ISBN: 9781450368919. DOI: 10.1145/3343055.3359718. URL: <https://doi.org/10.1145/3343055.3359718>.
- [190] Horst W. J. Rittel and Melvin M. Webber. “Dilemmas in a general theory of planning.” In: *Policy Sciences* 4.2 (June 1973), pp. 155–169. ISSN: 1573-0891. DOI: 10.1007/BF01405730. URL: <https://doi.org/10.1007/BF01405730>.
- [191] Kostia Robert, Dingyun Zhu, Weidong Huang, Leila Alem, and Tom Gedeon. “MobileHelper: Remote Guiding Using Smart Mobile Devices, Hand Gestures and Augmented Reality.” In: *SIGGRAPH Asia 2013 Symposium on Mobile Graphics and Interactive Applications*. SA ’13. Hong Kong, Hong Kong: Association for Computing Machinery, 2013. ISBN: 9781450326339. DOI: 10.1145/2543651.2543664. URL: <https://doi.org/10.1145/2543651.2543664>.
- [192] George Robertson, Mary Czerwinski, and Maarten Dantzich. “Immersion in Desktop Virtual Reality.” In: Jan. 1997, pp. 11–19. DOI: 10.1145/263407.263409.
- [193] Yvonne Rogers. *HCI Theory: Classical, Modern, and Contemporary*. Synthesis Lectures on Human-Centered Informatics. Morgan & Claypool Publishers, 2012.
- [194] 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 30th Annual ACM Symposium on User Interface Software and Technology*. UIST ’17. Québec City, QC, Canada: Association for Computing Machinery, 2017, 787–795. ISBN: 9781450349819. DOI: 10.1145/3126594.3126638. URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3126594.3126638>.
- [195] Joan Sol Roo and Martin Hachet. “Towards a hybrid space combining Spatial Augmented Reality and virtual reality.” In: *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. 2017, pp. 195–198. DOI: 10.1109/3DUI.2017.7893339.
- [196] Stéphane Safin, Christelle Boulanger, and Leclercq Pierre. “Premières évaluations d’un bureau virtuel pour un processus de conception augmenté.” In: Jan. 2005, pp. 107–114. DOI: 10.1145/1148550.1148564.

- [197] Kyung Sakong and Tek-jin Nam. "Supporting Telepresence by Visual and Physical Cues in Distributed 3D Collaborative Design Environments." In: *CHI '06 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '06. Montréal, Québec, Canada: Association for Computing Machinery, 2006, 1283–1288. ISBN: 1595932984. DOI: [10.1145/1125451.1125690](https://doi.org/10.1145/1125451.1125690). URL: <https://doi.org/10.1145/1125451.1125690>.
- [198] Patrick Salamin, Daniel Thalmann, and Frédéric Vexo. "The Benefits of Third-Person Perspective in Virtual and Augmented Reality?" In: *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*. VRST '06. Limassol, Cyprus: Association for Computing Machinery, 2006, 27–30. ISBN: 1595933212. DOI: [10.1145/1180495.1180502](https://doi-org.ins2i.bib.cnrs.fr/10.1145/1180495.1180502). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/1180495.1180502>.
- [199] Ugo Sangiorgi. "Addressing Multi-Platform Collaborative Sketching." In: *Proceedings of the 4th ACM SIGCHI Symposium on Engineering Interactive Computing Systems*. EICS '12. Copenhagen, Denmark: Association for Computing Machinery, 2012, 309–312. ISBN: 9781450311687. DOI: [10.1145/2305484.2305538](https://doi-org.ins2i.bib.cnrs.fr/10.1145/2305484.2305538). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/2305484.2305538>.
- [200] Sheree May Saßmannshausen, Jörg Radtke, Nino Bohn, Hassan Hussein, Dave Randall, and Volkmar Pipek. "Citizen-Centered Design in Urban Planning: How Augmented Reality Can Be Used in Citizen Participation Processes." In: *Designing Interactive Systems Conference 2021*. New York, NY, USA: Association for Computing Machinery, 2021, 250–265. ISBN: 9781450384766. URL: <https://doi.org/10.1145/3461778.3462130>.
- [201] Wendy A. Schafer and Doug A. Bowman. "Integrating 2D and 3D Views for Spatial Collaboration." In: *Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work*. GROUP '05. Sanibel Island, Florida, USA: Association for Computing Machinery, 2005, 41–50. ISBN: 1595932232. DOI: [10.1145/1099203.1099210](https://doi-org.ins2i.bib.cnrs.fr/10.1145/1099203.1099210). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/1099203.1099210>.
- [202] Jonas Schjerlund, Kasper Hornbæk, and Joanna Bergström. "OVRlap: Perceiving Multiple Locations Simultaneously to Improve Interaction in VR." In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. New Orleans, LA, USA: Association for Computing Machinery, 2022. ISBN: 9781450391573. DOI: [10.1145/3491102.3501873](https://doi.org/10.1145/3491102.3501873). URL: <https://doi.org/10.1145/3491102.3501873>.
- [203] Kjeld Schmidt. "The Problem with 'Awareness': Introductory Remarks on 'Awareness in CSCW'." In: *Computer Supported*

- Cooperative Work* 11 (Sept. 2002), pp. 285–298. DOI: [10.1023/A:1021272909573](https://doi.org/10.1023/A:1021272909573).
- [204] Michael F. Schober. “Speakers, addressees, and frames of reference: Whose effort is minimized in conversations about locations?” In: *Discourse Processes* 20.2 (1995), pp. 219–247. DOI: [10.1080/01638539509544939](https://doi.org/10.1080/01638539509544939). eprint: <https://www.tandfonline.com/doi/pdf/10.1080/01638539509544939>. URL: <https://www.tandfonline.com/doi/abs/10.1080/01638539509544939>.
- [205] Donald A Schön. *The reflective practitioner: How professionals think in action*. Routledge, 2017.
- [206] Stacey D. Scott, M. Sheelagh T. Carpendale, and Kori Inkpen. “Territoriality in Collaborative Tabletop Workspaces.” In: *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*. CSCW ’04. Chicago, Illinois, USA: Association for Computing Machinery, 2004, 294–303. ISBN: 1581138105. DOI: [10.1145/1031607.1031655](https://doi.org/10.1145/1031607.1031655). URL: <https://doi.org/10.1145/1031607.1031655>.
- [207] R. N. Shepard and J. Metzler. “Mental Rotation of Three-Dimensional Objects.” In: *Science* 171.3972 (Feb. 1971), pp. 701–703. DOI: [10.1126/science.171.3972.701](https://doi.org/10.1126/science.171.3972.701). URL: <https://doi.org/10.1126/science.171.3972.701>.
- [208] Joon Gi Shin, Gary Ng, and Daniel Saakes. “Couples Designing Their Living Room Together: A Study with Collaborative Handheld Augmented Reality.” In: *Proceedings of the 9th Augmented Human International Conference*. AH ’18. Seoul, Republic of Korea: Association for Computing Machinery, 2018. ISBN: 9781450354158. DOI: [10.1145/3174910.3174930](https://doi.org/10.1145/3174910.3174930). URL: <https://doi.org/10.1145/3174910.3174930>.
- [209] Qing Shuai, Chen Geng, Qi Fang, Sida Peng, Wenhao Shen, Xiaowei Zhou, and Hujun Bao. “Novel View Synthesis of Human Interactions from Sparse Multi-View Videos.” In: *ACM SIGGRAPH 2022 Conference Proceedings*. SIGGRAPH ’22. Vancouver, BC, Canada: Association for Computing Machinery, 2022. ISBN: 9781450393379. DOI: [10.1145/3528233.3530704](https://doi.org/10.1145/3528233.3530704). URL: <https://doi.org/10.1145/3528233.3530704>.
- [210] Karan Singh and Ravin Balakrishnan. “Visualizing 3D Scenes Using Non-Linear Projections and Data Mining of Previous Camera Movements.” In: *Proceedings of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*. AFRIGRAPH ’04. Stellenbosch, South Africa: Association for Computing Machinery, 2004, 41–48. ISBN: 1581138636. DOI: [10.1145/1029949.1029956](https://doi.org/10.1145/1029949.1029956). URL: <https://doi.org/10.1145/1029949.1029956>.

- [211] John G. Smith, David S. Miller, and Michael J. Muller. "TelePictive Groupware for Collaborative GUI Design." In: *Posters and Short Talks of the 1992 SIGCHI Conference on Human Factors in Computing Systems*. CHI '92. Monterey, California: Association for Computing Machinery, 1992, p. 41. ISBN: 9781450378048. DOI: [10.1145/1125021.1125059](https://doi.org/10.1145/1125021.1125059). URL: <https://doi.org/10.1145/1125021.1125059>.
- [212] Randall B. Smith. "What You See is What I Think You See." In: *SIGCUE Outlook* 21.3 (1992), 18–23. ISSN: 0163-5735. DOI: [10.1145/130893.130898](https://doi.org/10.1145/130893.130898). URL: <https://doi.org/10.1145/130893.130898>.
- [213] Rajinder S. Sodhi, Brett R. Jones, David Forsyth, Brian P. Bailey, and Giuliano Maciocci. "BeThere: 3D Mobile Collaboration with Spatial Input." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2013, 179–188. ISBN: 9781450318990. URL: <https://doi.org/10.1145/2470654.2470679>.
- [214] Edward W Soja. "The Political Organization of Space, Resource Paper No. 8." In: (1971).
- [215] Kihoon Son, Hwiwon Chun, Sojin Park, and Kyung Hoon Hyun. "C-Space: An Interactive Prototyping Platform for Collaborative Spatial Design Exploration." In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Honolulu, HI, USA: Association for Computing Machinery, 2020, 1–13. ISBN: 9781450367080. DOI: [10.1145/3313831.3376452](https://doi.org/10.1145/3313831.3376452). URL: <https://doi.org/10.1145/3313831.3376452>.
- [216] Shuang Song and Alice M. Agogino. "Insights on Designers' Sketching Activities in New Product Design Teams." In: 2004.
- [217] Maurício Sousa, Rafael Kufner dos Anjos, Daniel Mendes, Mark Billinghurst, and Joaquim Jorge. "Warping Deixis: Distorting Gestures to Enhance Collaboration." In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland UK: Association for Computing Machinery, 2019, 1–12. ISBN: 9781450359702. DOI: [10.1145/3290605.3300838](https://doi.org/10.1145/3290605.3300838). URL: <https://doi.org/10.1145/3290605.3300838>.
- [218] Maximilian Speicher, Brian D. Hall, Ao Yu, Bowen Zhang, Haihua Zhang, Janet Nebeling, and Michael Nebeling. "XD-AR: Challenges and Opportunities in Cross-Device Augmented Reality Application Development." In: *Proc. ACM Hum.-Comput. Interact.* 2.EICS (2018). DOI: [10.1145/3229089](https://doi.org/10.1145/3229089). URL: <https://doi.org/10.1145/3229089>.

- [219] Misha Sra, Sergio Garrido-Jurado, Chris Schmandt, and Pattie Maes. "Procedurally Generated Virtual Reality from 3D Reconstructed Physical Space." In: *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*. VRST '16. Munich, Germany: Association for Computing Machinery, 2016, 191–200. ISBN: 9781450344913. DOI: [10.1145/2993369.2993372](https://doi.org/10.1145/2993369.2993372). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/2993369.2993372>.
- [220] Aaron Stafford, Wayne Piekarski, and Bruce H. Thomas. "Implementation of god-like interaction techniques for supporting collaboration between outdoor AR and indoor tabletop users." In: *2006 IEEE/ACM International Symposium on Mixed and Augmented Reality*. 2006, pp. 165–172. DOI: [10.1109/ISMAR.2006.297809](https://doi.org/10.1109/ISMAR.2006.297809).
- [221] Joachim Stempfle and Petra Badke-Schaub. "Thinking in design teams - an analysis of team communication." In: *Design Studies* 23.5 (2002), pp. 473–496. ISSN: 0142-694X. DOI: [https://doi.org/10.1016/S0142-694X\(02\)00004-2](https://doi.org/10.1016/S0142-694X(02)00004-2). URL: <https://www.sciencedirect.com/science/article/pii/S0142694X02000042>.
- [222] Mengu Sukan, Steven Feiner, Barbara Tversky, and Semih Energin. "Quick Viewpoint Switching for Manipulating Virtual Objects in Hand-Held Augmented Reality Using Stored Snapshots." In: *Proceedings of the 2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. ISMAR '12. USA: IEEE Computer Society, 2012, 217–226. ISBN: 9781467346603. DOI: [10.1109/ISMAR.2012.6402560](https://doi.org/10.1109/ISMAR.2012.6402560). URL: <https://doi.org/10.1109/ISMAR.2012.6402560>.
- [223] Hongling Sun, Yue Liu, Zhenliang Zhang, Xiaoxu Liu, and Yongtian Wang. "Employing Different Viewpoints for Remote Guidance in a Collaborative Augmented Environment." In: *Proceedings of the Sixth International Symposium of Chinese CHI*. ChineseCHI '18. Montreal, QC, Canada: Association for Computing Machinery, 2018, 64–70. ISBN: 9781450365086. DOI: [10.1145/3202667.3202676](https://doi.org/10.1145/3202667.3202676). URL: <https://doi.org/10.1145/3202667.3202676>.
- [224] Ivan Sutherland. "The ultimate display." In: (1965).
- [225] Z. Szalavári, D. Schmalstieg, A. Fuhrmann, and M. Gervautz. "'Studierstube': An environment for collaboration in augmented reality." In: *Virtual Reality* 3.1 (Mar. 1998), pp. 37–48. ISSN: 1434-9957. DOI: [10.1007/BF01409796](https://doi.org/10.1007/BF01409796). URL: <https://doi.org/10.1007/BF01409796>.

- [226] Zsolt Szalavári and Michael Gervautz. "The Personal Interaction Panel - a Two-Handed Interface for Augmented Reality." In: *Comput. Graph. Forum* 16 (Sept. 1997), pp. 335–346. DOI: [10.1111/1467-8659.16.3conferenceissue.35](https://doi.org/10.1111/1467-8659.16.3conferenceissue.35).
- [227] Matthew Tait and Mark Billingham. "The Effect of View Independence in a Collaborative AR System." In: *Computer Supported Cooperative Work (CSCW)* 24 (Aug. 2015). DOI: [10.1007/s10606-015-9231-8](https://doi.org/10.1007/s10606-015-9231-8).
- [228] Chiew Seng Sean Tan, Kris Luyten, Jan Van Den Bergh, Johannes Schöning, and Karin Coninx. "The Role of Physiological Cues during Remote Collaboration." In: *Presence: Teleoperators and Virtual Environments* 23.1 (Feb. 2014), pp. 90–107. DOI: [10.1162/PRES_a_00168](https://doi.org/10.1162/PRES_a_00168). eprint: https://direct.mit.edu/pvar/article-pdf/23/1/90/1625375/pres_a_00168.pdf. URL: https://doi.org/10.1162/PRES_a_00168.
- [229] Arthur Tang, Charles Owen, Frank Biocca, and Weimin Mou. "Experimental Evaluation of Augmented Reality in Object Assembly Task." In: *Proceedings of the 1st International Symposium on Mixed and Augmented Reality*. ISMAR '02. USA: IEEE Computer Society, 2002, p. 265. ISBN: 0769517811.
- [230] John C. Tang. "Findings from Observational Studies of Collaborative Work." In: *Int. J. Man-Mach. Stud.* 34.2 (1991), 143–160. ISSN: 0020-7373. DOI: [10.1016/0020-7373\(91\)90039-A](https://doi.org/10.1016/0020-7373(91)90039-A). URL: [https://doi.org/10.1016/0020-7373\(91\)90039-A](https://doi.org/10.1016/0020-7373(91)90039-A).
- [231] 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: Association for Computing Machinery, 1990, 313–320. ISBN: 0201509326. DOI: [10.1145/97243.97302](https://doi.org/10.1145/97243.97302). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/97243.97302>.
- [232] Markus Tatzgern, Raphael Grasset, Denis Kalkofen, and Dieter Schmalstieg. "Transitional Augmented Reality navigation for live captured scenes." In: *2014 IEEE Virtual Reality (VR)*. 2014, pp. 21–26. DOI: [10.1109/VR.2014.6802045](https://doi.org/10.1109/VR.2014.6802045).
- [233] Ralph B. Taylor. *Human Territorial Functioning: An Empirical, Evolutionary Perspective on Individual and Small Group Territorial Cognitions, Behaviors, and Consequences*. Environment and Behavior. Cambridge University Press, 1988. DOI: [10.1017/CB09780511571237](https://doi.org/10.1017/CB09780511571237).
- [234] Franco Tecchia, Leila Alem, and Weidong Huang. "3D Helping Hands: A Gesture Based MR System for Remote Collaboration." In: *Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in*

- Industry*. VRCAI '12. Singapore, Singapore: Association for Computing Machinery, 2012, 323–328. ISBN: 9781450318259. DOI: [10.1145/2407516.2407590](https://doi.org/10.1145/2407516.2407590). URL: <https://doi.org/10.1145/2407516.2407590>.
- [235] 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 2019 CHI Conference on Human Factors in Computing Systems*. CHI '19. Glasgow, Scotland Uk: Association for Computing Machinery, 2019, 1–14. ISBN: 9781450359702. DOI: [10.1145/3290605.3300431](https://doi.org/10.1145/3290605.3300431). URL: <https://doi.org/10.1145/3290605.3300431>.
- [236] Theophilus Teo, Gun A. Lee, Mark Billinghurst, and Matt Adcock. “Hand Gestures and Visual Annotation in Live 360 Panorama-Based Mixed Reality Remote Collaboration.” In: *Proceedings of the 30th Australian Conference on Computer-Human Interaction*. OzCHI '18. Melbourne, Australia: Association for Computing Machinery, 2018, 406–410. ISBN: 9781450361880. DOI: [10.1145/3292147.3292200](https://doi.org/10.1145/3292147.3292200). URL: <https://doi.org/10.1145/3292147.3292200>.
- [237] Santawat Thanyadit and Ting-Chuen Pong. “User Interface Applications in Desktop VR Using a Mirror Metaphor.” In: *SIGGRAPH Asia 2017 Posters*. SA '17. Bangkok, Thailand: Association for Computing Machinery, 2017. ISBN: 9781450354059. DOI: [10.1145/3145690.3145716](https://doi.org/10.1145/3145690.3145716). URL: <https://doi.org/10.1145/3145690.3145716>.
- [238] Balasaravanan Thoravi Kumaravel, Fraser Anderson, George Fitzmaurice, Bjoern Hartmann, and Tovi Grossman. “Loki: Facilitating Remote Instruction of Physical Tasks Using Bi-Directional Mixed-Reality Telepresence.” In: *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. UIST '19. New Orleans, LA, USA: Association for Computing Machinery, 2019, 161–174. ISBN: 9781450368162. DOI: [10.1145/3332165.3347872](https://doi-org.ins2i.bib.cnrs.fr/10.1145/3332165.3347872). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3332165.3347872>.
- [239] Balasaravanan Thoravi Kumaravel, Cuong Nguyen, Stephen DiVerdi, and Bjoern Hartmann. “TransceiVR: Bridging Asymmetrical Communication Between VR Users and External Collaborators.” In: *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. UIST '20. Virtual Event, USA: Association for Computing Machinery, 2020, 182–195. ISBN: 9781450375146. DOI: [10.1145/3379337.3415827](https://doi.org/10.1145/3379337.3415827). URL: <https://doi.org/10.1145/3379337.3415827>.
- [240] Walter F Tichy. “RCS—A system for version control.” In: *Software: Practice and Experience* 15.7 (1985), pp. 637–654.

- [241] Baris Unver, Sarah D'Angelo, Matthew Miller, John C. Tang, Gina Venolia, and Kori Inkpen. "Hands-Free Remote Collaboration Over Video: Exploring Viewer and Streamer Reactions." In: *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*. ISS '18. Tokyo, Japan: Association for Computing Machinery, 2018, 85–95. ISBN: 9781450356947. DOI: [10.1145/3279778.3279803](https://doi.org/10.1145/3279778.3279803). URL: <https://doi.org/10.1145/3279778.3279803>.
- [242] Eswara Rao Velamkayala, Manuel V. Zambrano, and Huiyang Li. "Effects of HoloLens in Collaboration: A Case in Navigation Tasks." In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 61.1 (2017), pp. 2110–2114. DOI: [10.1177/1541931213602009](https://doi.org/10.1177/1541931213602009). eprint: <https://doi.org/10.1177/1541931213602009>. URL: <https://doi.org/10.1177/1541931213602009>.
- [243] Alonso H. Vera, Thomas Kvan, Robert L. West, and Simon Lai. "Expertise, Collaboration and Bandwidth." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '98. Los Angeles, California, USA: ACM Press/Addison-Wesley Publishing Co., 1998, 503–510. ISBN: 0201309874. DOI: [10.1145/274644.274712](https://doi-org.ins2i.bib.cnrs.fr/10.1145/274644.274712). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/274644.274712>.
- [244] John Viega, Matthew J. Conway, George Williams, and Randy Pausch. "3D Magic Lenses." In: *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology*. UIST '96. Seattle, Washington, USA: Association for Computing Machinery, 1996, 51–58. ISBN: 0897917987. DOI: [10.1145/237091.237098](https://doi.org/10.1145/237091.237098). URL: <https://doi.org/10.1145/237091.237098>.
- [245] Philipp Wacker, Adrian Wagner, Simon Voelker, and Jan Borchers. "Physical Guides: An Analysis of 3D Sketching Performance on Physical Objects in Augmented Reality." In: *Proceedings of the Symposium on Spatial User Interaction*. SUI '18. Berlin, Germany: Association for Computing Machinery, 2018, 25–35. ISBN: 9781450357081. DOI: [10.1145/3267782.3267788](https://doi.org/10.1145/3267782.3267788). URL: <https://doi.org/10.1145/3267782.3267788>.
- [246] Flávio R. Wagner and Arnaldo H. Viegas de Lima. "Design Version Management in the GARDEN Framework." In: *Proceedings of the 28th ACM/IEEE Design Automation Conference*. DAC '91. San Francisco, California, USA: Association for Computing Machinery, 1991, 704–710. ISBN: 0897913957. DOI: [10.1145/127601.127755](https://doi.org/10.1145/127601.127755). URL: <https://doi.org/10.1145/127601.127755>.
- [247] 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: *Proc. ACM Hum.-*

- Comput. Interact.* 4.CSCW2 (2020). DOI: [10.1145/3415243](https://doi-org.ins2i.bib.cnrs.fr/10.1145/3415243). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3415243>.
- [248] 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 33rd Annual ACM Symposium on User Interface Software and Technology*. UIST '20. Virtual Event, USA: Association for Computing Machinery, 2020, 805–817. ISBN: 9781450375146. DOI: [10.1145/3379337.3415868](https://doi-org.ins2i.bib.cnrs.fr/10.1145/3379337.3415868). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3379337.3415868>.
- [249] Isaac Wang, Pradyumna Narayana, Dhruva Patil, Gururaj Mulla, Rahul Bangar, Bruce Draper, Ross Beveridge, and Jaime Ruiz. "Exploring the Use of Gesture in Collaborative Tasks." In: *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. CHI EA '17. Denver, Colorado, USA: Association for Computing Machinery, 2017, 2990–2997. ISBN: 9781450346566. DOI: [10.1145/3027063.3053239](https://doi.org/10.1145/3027063.3053239). URL: <https://doi.org/10.1145/3027063.3053239>.
- [250] John Wang and Edwin Olson. "AprilTag 2: Efficient and robust fiducial detection." In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2016.
- [251] Peng Wang, Shusheng Zhang, Mark Billingham, Xiaoliang Bai, Weiping He, Shuxia Wang, Mengmeng Sun, and Xu Zhang. "A comprehensive survey of AR/MR-based co-design in manufacturing." In: *Engineering with Computers* 36.4 (Oct. 2020), pp. 1715–1738. ISSN: 1435-5663. DOI: [10.1007/s00366-019-00792-3](https://doi.org/10.1007/s00366-019-00792-3). URL: <https://doi.org/10.1007/s00366-019-00792-3>.
- [252] Thomas Wells and Steven Houben. "CollabAR - Investigating the Mediating Role of Mobile AR Interfaces on Co-Located Group Collaboration." In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Honolulu, HI, USA: Association for Computing Machinery, 2020, 1–13. ISBN: 9781450367080. DOI: [10.1145/3313831.3376541](https://doi.org/10.1145/3313831.3376541). URL: <https://doi.org/10.1145/3313831.3376541>.
- [253] Michael Wessely, Theophanis Tsandilas, and Wendy E. Mackay. "Shape-Aware Material: Interactive Fabrication with ShapeMe." In: *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. UIST '18. Berlin, Germany: Association for Computing Machinery, 2018, 127–139. ISBN: 9781450359481. DOI: [10.1145/3242587.3242619](https://doi-org.ins2i.bib.cnrs.fr/10.1145/3242587.3242619). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3242587.3242619>.

- [254] Michael Wittkämper, Irma Lindt, Wolfgang Broll, Jan Ohlenburg, Jan Herling, and Sabiha Ghellal. “Exploring Augmented Live Video Streams for Remote Participation.” In: *CHI '07 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '07. San Jose, CA, USA: Association for Computing Machinery, 2007, 1881–1886. ISBN: 9781595936424. DOI: [10.1145/1240866.1240915](https://doi.org/10.1145/1240866.1240915). URL: <https://doi.org/10.1145/1240866.1240915>.
- [255] Jacob O. Wobbrock and Julie A. Kientz. “Research Contributions in Human-Computer Interaction.” In: *Interactions* 23.3 (2016), 38–44. ISSN: 1072-5520. DOI: [10.1145/2907069](https://doi.org/10.1145/2907069). URL: <https://doi.org/10.1145/2907069>.
- [256] Dustin Wüest, Norbert Seyff, and Martin Glinz. “FLEXISKETCH TEAM: Collaborative Sketching and Notation Creation on the Fly.” In: *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering*. Vol. 2. 2015, pp. 685–688. DOI: [10.1109/ICSE.2015.223](https://doi.org/10.1109/ICSE.2015.223).
- [257] Haijun Xia, Sebastian Herscher, Ken Perlin, and Daniel Wigdor. “Spacetime: Enabling Fluid Individual and Collaborative Editing in Virtual Reality.” In: *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. UIST '18. Berlin, Germany: Association for Computing Machinery, 2018, 853–866. ISBN: 9781450359481. DOI: [10.1145/3242587.3242597](https://doi.org/10.1145/3242587.3242597). URL: <https://doi.org/10.1145/3242587.3242597>.
- [258] Longqi Yang et al. “The effects of remote work on collaboration among information workers.” In: *Nature Human Behaviour* (2021). DOI: [10.1038/s41562-021-01196-4](https://doi.org/10.1038/s41562-021-01196-4). URL: <https://doi.org/10.1038/s41562-021-01196-4>.
- [259] Maria Yang and Jorge Cham. “An Analysis of Sketching Skill and Its Role in Early Stage Engineering Design.” In: May 2007.
- [260] Alper Yilmaz, Omar Javed, and Mubarak Shah. “Object Tracking: A Survey.” In: *ACM Comput. Surv.* 38.4 (2006), 13–es. ISSN: 0360-0300. DOI: [10.1145/1177352.1177355](https://doi.org/10.1145/1177352.1177355). URL: <https://doi.org/10.1145/1177352.1177355>.
- [261] 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](https://doi.org/10.1109/TVCG.2022.3150520).
- [262] Yuchong Zhang, Adam Nowak, Andrzej Romanowski, and Morten Fjeld. “On-Site or Remote Working?: An Initial Solution on How COVID-19 Pandemic May Impact Augmented Reality Users.” In: *Proceedings of the 2022 International Conference on Advanced Visual Interfaces*. AVI 2022. Frascati, Rome, Italy: Association for Computing Machinery, 2022. ISBN: 9781450397193.

- DOI: [10.1145/3531073.3534490](https://doi-org.ins2i.bib.cnrs.fr/10.1145/3531073.3534490). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/3531073.3534490>.
- [263] Zhenpeng Zhao, Sriram Karthik Badam, Senthil Chandrasegaran, Deok Gun Park, Niklas L.E. Elmqvist, Lorraine Kisselburgh, and Karthik Ramani. "SkWiki: A Multimedia Sketching System for Collaborative Creativity." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '14. Toronto, Ontario, Canada: Association for Computing Machinery, 2014, 1235–1244. ISBN: 9781450324731. DOI: [10.1145/2556288.2557394](https://doi-org.ins2i.bib.cnrs.fr/10.1145/2556288.2557394). URL: <https://doi-org.ins2i.bib.cnrs.fr/10.1145/2556288.2557394>.
- [264] Jakob Zillner, Erick Mendez, and Daniel Wagner. "Augmented Reality Remote Collaboration with Dense Reconstruction." In: *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 2018, pp. 38–39. DOI: [10.1109/ISMAR-Adjunct.2018.00028](https://doi-org.ins2i.bib.cnrs.fr/10.1109/ISMAR-Adjunct.2018.00028).
- [265] Cinemachine. Suite of modules for operating the Unity camera <https://unity.com/fr/unity/features/editor/art-and-design/cinemachine>. Retrieved July 28th 2021. 2021.
- [266] Hololens 2. Microsoft Hololens 2 AR Headset. <https://www.microsoft.com/fr-fr/hololens/hardware>. Retrieved January 22th 2021. 2021.
- [267] MRTK. Mixed Reality Toolkit (MRTK) provides a set of components and features to accelerate cross-platform MR app development in Unity. <https://github.com/microsoft/MixedRealityToolkit-Unity>. Retrieved October 17th 2022. 2022.
- [268] Microsoft HoloLens Application. Microsoft Hololens 2 Desktop Application <https://www.microsoft.com/en-us/p/microsoft-hololens/9nblggh4qwnx>. Retrieved July 26th 2021. 2021.
- [269] MixedReality-WebRTC. Microsoft Mixed Reality WebRTC libraries. <https://github.com/microsoft/MixedReality-WebRTC>. Retrieved July 26th 2021. 2021.
- [270] Sketchboard. Sketchboard's virtual whiteboard solution. <https://sketchboard.io/>. Retrieved October 10th 2022. 2022.
- [271] UN/DESA Policy Brief 92: Leveraging digital technologies for social inclusion. United Nation, Department of Economic and Social Affairs, Economic Analysis. <https://www.un.org/development/desa/dpad/publication/un-des-a-policy-brief-92-leveraging-digital-technologies-for-social-inclusion/>. Retrieved November 26th 2021. 2021.

- [272] UN/DESA Policy Brief 92: Leveraging digital technologies for social inclusion. United Nation, Department of Economic and Social Affairs, Economic Analysis. <https://www.un.org/development/desa/dpad/publication/un-desa-policy-brief-92-leveraging-digital-technologies-for-social-inclusion/>. Retrieved November 26th 2021. 2021.
- [273] UNet. Unity Multiplayer and Networking. <https://docs.unity3d.com/Manual/UNet.html>. Retrieved January 22th 2021. 2021.
- [274] Unity. Unity Technologies. <https://unity.com/fr>. Retrieved January 22th 2021. 2021.
- [275] WebRTC. WebRTC Real time communication for the Web. <https://webrtc.org/>. Retrieved January 22th 2021. 2021.