CHALMERS

Designing to Support Workspace Awareness in Remote Collaboration using 2D Interactive Surfaces

KHANH-DUY LE

Department of Computer Science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

Thesis for the degree of Doctor of Philosophy

Designing to Support Workspace Awareness in Remote Collaboration using 2D Interactive Surfaces

Khanh-Duy Le

Department of Computer Science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2019

Designing to Support Workspace Awareness in Remote Collaboration using 2D Interactive Surfaces
KHANH-DUY LE
ISBN 978-91-7905-199-0

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 4666
ISSN 0346-718X

Department of Computer Science and Engineering
Chalmers University of Technology
SE–412 96 Göteborg, Sweden
Telephone + 46 (0) 31 – 772 1000

Cover: VXSlate - envisioned concept synthesizing pieces of work in the thesis, see Figure 6.1 on page 59. The cover image is designed by the author.

Printed by Chalmers Reproservice
Göteborg, Sweden 2019

to my loved ones

Designing to Support Workspace Awareness in
Remote Collaboration using 2D Interactive Surfaces
Thesis for the Degree of Doctor of Philosophy
KHANH-DUY LE
Department of Computer Science and Engineering
Chalmers University of Technology

Abstract

Increasing distribution of the global workforce are leading to collaborative work among remote coworkers. The emergence of such remote collaboration is essentially supported by technology advancements of screen-based devices ranging from tablet or laptop to large displays. However, these devices, especially personal and mobile computers, still suffer from certain limitations caused by their form factors, that hinder supporting workspace awareness through non-verbal communication such as bodily gestures or gaze. This thesis thus aims to design novel interfaces and interaction techniques to improve remote coworkers' workspace awareness through such non-verbal cues using 2D interactive surfaces.

The thesis starts off by exploring how visual cues support workspace awareness in facilitated brainstorming of hybrid teams of co-located and remote coworkers. Based on insights from this exploration, the thesis introduces three interfaces for mobile devices that help users maintain and convey their workspace awareness with their coworkers. The first interface is a virtual environment that allows a remote person to effectively maintain his/her awareness of his/her co-located collaborators' activities while interacting with the shared workspace. To help a person better express his/her hand gestures in remote collaboration using a mobile device, the second interface presents a lightweight add-on for capturing hand images on and above the device's screen; and overlaying them on collaborators' device to improve their workspace awareness. The third interface strategically leverages the entire screen space of a conventional laptop to better convey a remote person's gaze to his/her co-located collaborators. Building on the top of these three interfaces, the thesis envisions an interface that supports a person using a mobile device to effectively collaborate with remote coworkers working with a large display.

Together, these interfaces demonstrate the possibilities to innovate on commodity devices to offer richer non-verbal communication and better support workspace awareness in remote collaboration.

Keywords: Remote collaboration, 2D interactive surfaces, gaze, hand gestures, virtual collaboration, mobile devices, telepresence

Acknowledgments

Looking back the path that I have been going on, I realize that I am really fortunate. By several coincidences and generous support from many people, I'm still childishly enjoying working with the things that I have been fascinated with for more than twenty five years. In this little piece of text, I will try to express my gratitude to as many of those people as possible. However, I believe that there are still ones, who at some point in my life had certain influence to help me become who I am today, not covered in this text. If that is the case, I hope that they will accept my sincere apology and understand that I really appreciate what they have done for me.

The very first person that I would like to thank is of course my supervisor — Prof. Morten Fjeld, who made this position possible, fully supported whatever I wanted to do, and tolerantly empathized with my nastiness at some points during this journey. I would also have never gone this far without the companionship of my co-advisor, Prof. Andreas Kunz, who always arranged his time to discuss ideas with me and provided timely advice for my decision making, regardless of our geographical distance and his dense schedule.

As usual, whenever I reach a milestone on my journey of human-computer interaction, I always recall Prof. Minh-Triet Tran , who enlightened me about the existence of this field, helped me realize what my passion was and simply turned my life towards a very different page. Also I will never forget the stepping-stone opportunities that Prof. Vu Duong gave me to be educated, work and develop myself in advanced European environments.

Even though I tried to travel to Vietnam as much as I could, I know that I never spent enough time being beside my parents. I feel really blessed for having Khanh-Vinh Le and Khanh-Trung Le as my brothers, whom I need to thank for taking care of our parents and simply for the moments we shared together, well, as brothers. Thanks for my sisters in law for being a part of my family and giving me new family members that I can visit and spend time with whenever I'm in Vietnam.

Thanks Tomasz Kosinski, Paweł W Woźniak, Alexandru Dancu, Mehmet Aydın Baytaş, Jesper Molin and Zlatko Franjcic for being my current and former labmates and all the discussions we had together as well as your generous support regardless of whether you are still in the lab or not. I also deeply appreciate Prof. Marco Fratarcageli's mentorship for my career, especially when I was in the valley points of the journey. For international collaborators especially Kening Zhu, Prof. Shengdong Zhao, Ignacio Avellino and Cédric Fleury, I can only say that I learned a lot from them and really appreciate their openness to collaborate with a novice researcher like me. Likewise, I also would like to express my appreciation to former lab affiliates such as Miguel Gargallo, Gabrielė Kasparavičiūtė and Adviye Ayça Ünlüer Çimen for their help in my research and career. I also really value the experiences and collaborations I had with the seniors in the division of Interaction Design. Additionally, I would like to thank all anonymous reviewers for their feedback and critiques on my submissions which really helped me sharpen my research skills, although I did not always agree with all of them. Also, thanks colleagues from the division of Software Engineering for bringing me along to their football matches, where I joyfully and failedly experimented with different amateur playing skills and tactics.

I would not have enjoyed my life in Gothenburg as much as I did without my Vietnamese friends and their families, who have been building a second family of mine here. I'm delightedly grateful that my friends in France; either Vietnamese, French or from other countries, still kept me in the loop of their lives although I have not been back there for a long time.

Finally, to my beloved My Phuong, thank you for being remotely beside me almost every single day, accompanying me through endless challenges and having chosen to stay with me until now regardless all the things you had to suffer. I know I'm still owing you the sunny days when we will be roaming around the breath-taking mountainous landscape of the Alps or peacefully enjoying a ride through the Tuscany countryside. I hope you understand that I have been always trying hard to realize that soon.

Khanh-Duy Le
Göteborg, June 2019

Preface

I guess my journey in the field of human computer interaction was ignited when I was still a kid living in a small village in Vietnam. Back then, I was always fascinated by the futuristic computer interfaces that I saw in Hollywood sci-fi movies and then trying to replicate them through my sketches on the yard of my house. However, even in my craziest dream, I had never imagined that my future would be as a person creating and exploring such things everyday. Of course, in the wildest thoughts, neither my family nor I would ever dare to believe that one day I could get this far, writing this manuscript for a doctoral degree.

Personally, I have never seen myself as a true scholar as I enjoy creating tangible artifacts rather than producing knowledge. Frankly, when I started this position, I just considered this as a chance to realize the futuristic interfaces that I had been hooking into. Deriving scientific knowledge from such creations was not considered rigorously at that beginning. Nevertheless, I gradually changed my mind, even though I have to confess that it was initially to satisfy submission requirements in the field. I would say that such requirements are really meaningful to me. While trying to learn about interfaces and interaction techniques from user-centered and somehow theoretical perspectives, I could figure out that the actual beauty of the futuristic interfaces that attracted me lay in their usability — how they might help improve users' performance in their tasks. Sometimes, it was also quite heartbreaking for me to realize that some concepts I used to fancy are actually not that useful to users. However, this painful experience was a necessary step for me towards creating futuristic values that are applicable in reality. Honestly, it has an important role in shaping my doctoral research that will be presented in this manuscript. My doctoral research is also the reflection of my current research mindset - supporting futuristic interactions using pragmatic interface designs, which I hope this thesis can adequately convey to readers.

In the end, although I have gotten myself acquainted with certain theoretical aspects in human-computer interaction, I would still position myself as a maker, a creator or, flashier, an inventor, who loves creating novel interfaces and ways of interaction that hopefully help people live and work more conveniently and joyfully. However, whatever I regard myself today, it all started from the little kid sketching futuristic concepts on the ground in a small village in Vietnam twenty five years ago. For that reason, I hope that what is presented in this thesis can make him satisfied with regard to his imagination.

Khanh-Duy Le

List of Publications

This thesis is based on the following appended papers:

Paper A. Khanh-Duy Le, Paweł W. Woźniak, Ali Alavi, Morten Fjeld and Andreas Kunz. *DigiMetaplan: Supporting Facilitated Brainstorming for Distributed Business Teams.* Accepted for the 18th International Conference on Mobile and Ubiquitous Multimedia, Pisa, Italy, 2019.

Author's contribution: I was the primary designer and developer of the concept. I was also the primary author of the text.

Paper B. Khanh-Duy Le, Morten Fjeld, Ali Alavi and Andreas Kunz. *Immersive environment for distributed creative collaboration.* Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology, Gothenburg, Sweden, 2017.

Author's contribution: I was the sole designer and developer of the concept. I also wrote up the technical description and the future work of the paper.

Paper C. Khanh-Duy Le, Kening Zhu, Morten Fjeld. *Mirrortablet: Exploring a low-cost mobile system for capturing unmediated hand gestures in remote collaboration.* Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia, Stuttgart, Germany, 2017.

Author's contribution: I was the main designer and the sole developer of the system. I co-designed the user study with Kening Zhu; conducted and analyzed the study with his advice. I was also the primary author of the text.

Paper D. Khanh-Duy Le, Ignacio Avellino, Cédric Fleury, Morten Fjeld. *GazeLens: Guiding Attention to Improve Gaze Interpretation in Hub-Satellite Collaboration.* Proceedings of 17th IFIP TC 13 International Conference on Human-Computer Interaction, Paphos, Cyprus, 2019.

Author's contribution: I was the primary designer and the sole developer of the system . I co-designed the user study and conducted the experimental analysis with advice from Ignacio Avellino and Cédric Fleury. I was also the primary author of the text.

Other relevant publications co-authored by Khanh-Duy Le:

Vinh-Tiep Nguyen, **Khanh-Duy Le**, Minh-Triet Tran, Morten Fjeld. *NowAndThen: a social network-based photo recommendation tool supporting reminiscence.* Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia, Rovaniemi, Finland, 2016.

Author's contribution: I was the primary designer of the concept and contributed around 70% of the paper's text.

Khanh-Duy Le, Mahsa Paknezhad, Paweł W Woźniak, Maryam Azh, Gabrielė Kasparavičiūtė, Morten Fjeld, Shengdong Zhao, Michael S Brown. *Towards Leaning Aware Interaction with Multitouch Tabletops.* Proceedings of the 9th Nordic Conference on Human-Computer Interaction, Gothenburg, Sweden, 2016.

Author's contribution: I was the primary contributor to designing, developing experimental tools and conducting the user study. I also rewrote the applications of the interaction technique to make it accepted in the resubmission venue.

Khanh-Duy Le, Kening Zhu, Tomasz Kosinski, Morten Fjeld, Maryam Azh, Shengdong Zhao. *Ubitile: A finger-worn i/o device for tabletop vibrotactile pattern authoring.* Proceedings of the 9th Nordic Conference on Human-Computer Interaction, Gothenburg, Sweden, 2016.

Author's contribution: I brainstormed the idea with other co-authors. In addition, I was the primary designer and developer of the prototype. Finally, I was the main contributor to the text.

Rodi Jolak, **Khanh-Duy Le**, Kaan Burak Sener, Michel R.V. Chaudron. *OctoBubbles: A multi-view interactive environment for concurrent visualization and synchronization of UML models and code.* 2018 IEEE 25th International Conference on Software Analysis, Evolution and Reengineering (SANER), Campobasso, Italy, 2018.

Author's contribution: I wrote up the user study design on how to evaluate the design from human-centered perspectives.

List of Acronyms

HCI	–	Human-Computer Interaction
CSCW	–	Computer-Supported Collaborative Work
MR	–	Mixed Reality
VR	–	Virtual-Reality
AR	–	Augmented-Reality
2D	–	Two-Dimensional
2.5D	–	Two-and-a-Half-Dimensional
3D	–	Three-Dimensional
NASA	–	National Aeronautics and Space Administration
NASA TLX	–	NASA Task Load Index
MERCO	–	Mediated Remote Collaboration
FOV	–	Field of View

List of Figures

Contents

II Appended Papers 85

Part I

Introductory Chapters

Chapter 1

Setting the Scene

1.1 Introduction

"I slip into my routine in Bali with ease: yoga at 7am, a hemp smoothie for breakfast, a raw-food lunch, then working on my computer with my clients in Austria and UK, who are several hours different from me, and slinking off at 8pm for dinner". That is how Anna Hart described her work life as a "digital nomad", living in Indonesia while remotely running her business with her geographically-dispersed clients, published in The Telegraph in 2015 [1]. Digital nomads like Anna, referring to people using telecommunication technologies to earn a living and, more generally, conduct their life in a nomadic manner [2], are a new type of human workforce, gradually becoming mainstream [3]. This style of work life hugely benefits from technology developments especially in personal computing and telecommunication. With a laptop or a tablet connected to the internet, they can communicate with their business partners, get contracts settled, accomplish several kinds of tasks and deliver different types of business products almost anywhere and anytime (e.g. while sitting on a train, a bus or an airplane or lying down on a beach). Due to its mobility and flexibility advantages, many even believe that digital nomadism and remote working will become the future of the human workforce.

Remote working and digital nomadism not only provide a comfortable work experience to workers but also financial benefits to employers. Effective remote working can help companies and corporates reduce business traveling expenses of employees for face-to-face meetings. It can also help cut off budgets for employees' relocation. Therefore, many companies and organizations are experimenting digital nomadism with their employees. Companies like GitHub, ePublishing, OnTheGoSystems, or Working Solutions have already turned to hiring digital nomads. Most recently, Chalmers University of Technology experimented with reducing the number of private offices and increasing open working spaces instead. This means that

[1] https://www.telegraph.co.uk/news/features/11597145/Living-and-working-in-paradise-the-rise-of-the-digital-nomad.html

[2] https://en.wikipedia.org/wiki/Digital_nomad

[3] https://www.forbes.com/sites/elainepofeldt/2018/08/30/digital-nomadism-goes-mainstream

employees were encouraged to work with more location flexibility and mobility. They can stay at home while finishing their job using their computer and just need to stop by the school if needed (e.g. department or division meetings). A distributed setting commonly found nowadays is hub-satellite, referring to scenarios where one or multiple nomadic individuals (referred to as satellite workers) remotely work with their co-located team (referred to as hub). Such team settings can be found in several contexts: a design team's meeting, a project planning session, a follow-up meeting of a doctoral student's committee or a meeting of consultants and clients on certain products, etc. This kind of team settings has also been documented as trendy in academic literature but is still perceived as particularly receiving inadequate support, especially for the satellite worker [81, 143, 88, 91].

Any significant changes in the way humans work usually began from disruptive technology developments. For example, the introduction of engines in the 18th century led to the first industrialization, resulting in a new class of human labour called factory workers, who used to be farmers or craftsmen but were now brought to factories and worked together on giant machines. Similarly, the emergence of digital nomadism and remote working today is driven by the rapid development of computing devices. Since the introduction of the first personal computer in the 1970s, we have witnessed significant transformations of computer interfaces. Computers nowadays do not solely refer to box-size desktop devices but also wall-size screens at public places such as airports, train stations and commercial complexes or handheld devices such as tablets, smartphones, and even wearable form factors such as smartwatches and rings. In offices, large interactive displays are gradually replacing whiteboards and flip-charts due to their viewability to multiple people in a room and flexible interactions with digital contents displayed on them. With powerful processing units combined with rich sensors and connectivity packaged in compact form factors, personal devices like laptops, tablets and even phones allow users to accomplish several tasks from finishing business deliverables to video-conferencing with a distant partner while being on the go. These technology advancements serve as a necessary condition for the human workforce to desire a shift in the way they work together, even in complicated tasks such as creative problem solving or technology support, from physically collocated to remote working, as the latter offers more flexibility, comfortability and economic benefits for dispersed teams and organizations.

However, we are not there yet where most of people can work remotely. According to a research conducted by the Global Business Travel Association, global business travel spending topped a record-breaking $1.2 trillion in 2015, reached $1.33 trillion in 2017 and is predicted to reach $1.6 trillion in 2020 [4]. Figure 1.1 shows the global business travel spendings from 2015 to 2017 [5]. Even though business trips are not only costly but also uncomfortable due to their intense and short duration, business travel spendings are still on the rise. Regardless of the adoption of various telecommunication technologies such as email, shared documents, and video conferencing systems, face-to-face meetings are still the king when it comes

[4] https://www.inc.com/peter-economy/7-surprising-tricks-for-savvy-business-travelers.html

[5] https://www.statista.com/statistics/612244/global-business-travel-spending/

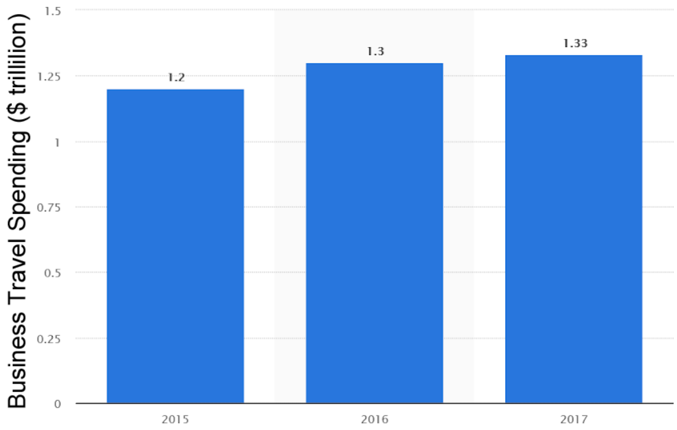

Figure 1.1: Global business travel spending from 2015 to 2017

to certain crucial collaborative tasks such as project milestone meeting, creative problem solving, and technology support. Although we have witnessed the rapidly increased internet bandwidth and the booming quantity of teleconference solutions, existing combinations of technology infrastructures and interface designs still cannot adequately resemble different characteristics of face-to-face collaboration. While current teleconference solutions can already allow users to hear and see their remote partners via audio and video channel, several subtle nonverbal communication cues cannot be successfully conveyed in their interfaces.

In “Distance matters” [108], one of the most impactful publications in the field of computer supported cooperative work (CSCW), Gary M. Olson and Judith S. Olson, suggested that in order for people to collaborate successfully, they need to have an adequate level of common ground. Common ground refers to the knowledge that the participants in a collaboration have in common, and they are aware that they have in common. The knowledge forming the participants’ common ground includes not only the understanding on the collaborative tasks, but also the comprehension on the participants’ status such as where they are and what they are doing with regard to the tasks. Therefore, to achieve a high level of common ground, the environment for collaboration must be able to support the participants to be adequately aware of various nonverbal cues of the others. Gary M. Olson and Judith S. Olson exemplified this with a meeting of a design team where all the team members were collocated. In such settings, a team member can be easily aware of a team member’s attention to some direction (e.g. pointing to drawings on a flip-chart, or gazing at someone or something in the room). Or a team member can also easily interpret the idea of another one by being aware of his/her gestures. Those aforementioned nonverbal cues are subtle but always exist in almost every face-to-face collaboration, making in-person communication intuitive and efficient.

Nevertheless, when one or more collaborators are not co-located, the situation can get more complicated. For example, with a design team scenario, when a collaborator is not in the same room, maintaining awareness to achieve high levels of common ground, requires conveying adequately the team’s space to the remote collaborator (e.g. the presence of people and artifacts as well as their locations in the room), and

Figure 1.2: Illustration envisioning outcomes of MERCO project

to support the remote collaborator's access to shared artifacts (i.e. artifacts involved in the collaboration session that all participants have the right to access such as flip-charts or physical prototypes). Likewise, conveying the remote collaborator's attention (e.g. pointing at shared artifacts or gazing at someone or something in the room) is also crucial to keep the collocated team aware of the counter part.

Although humans have been witnessing the rocketing development of computers and their related infrastructure to make remote collaboration more efficient and prevalent, there are still several technological issues to address, including designing suitable computer interfaces to support intuitive interaction and effective remote communication. While novel computing devices have been in research and development, promising non-precedential affordances, interaction and experiences, 2D interactive surfaces such as large displays, laptops, tablets or even smartphones will long still be considered as commodity devices for a majority of people. Hence, designing to support awareness in communication and interaction using such devices plays an important role in fostering remote collaboration. Although current interactive surfaces can support rich interaction modalities, there are still certain limitations caused by their form factor that can make conveying nonverbal communication cues challenging in remote collaboration. As it is not practical for nomadic workers to carry devices that are typically used in meeting rooms to capture and present non-verbal communication, it is thus essential to provide them more compact and lightweight solutions that allow them to achieve similar results with a high level of mobility. Throughout the works included in this thesis, my contributions demonstrate approaches to leverage and complement the form factor and affordances of off-the-shelf 2D interactive interfaces to facilitate users' awareness on their collaboration partner. My approaches rely either solely on software-based solutions or minimal hardware instrumentation on off-the-shelf devices, promising their applicability to reality.

1.2 MERCO project

Moving along with the global trend of fostering remote working in business, several industrial partners of our lab expressed their interest in researching and developing novel solutions for remote collaboration. This interest led to a collaborative research project called MERCO (**M**ediated **E**ffective **R**emote **Co**llaboration) among seven partner organizations from two countries, starting in 2014. The project partners included two academic research institutions (Chalmers University of Technology (Sweden) and ETH Zurich (Switzerland)) and five industrial organizations (Ericsson AB, SEMCON AB, Touchtech AB (Sweden), AVS Systeme AG, and Intelliconcept AG (Switzerland)). It is worth noting that the industrial partners had different business interests and expectations in this project. Ericsson AB was looking for solutions to support collaborative creative problem-solving activities in the company such as brainstorming in remote settings due to the distribution of the company's offices all over the world. SEMCON AB, as an international industrial-design consulting firm, expected to have cutting-edge remote collaboration to raise its appeal to customers and clients worldwide as well as using it for internal processes. Likewise, Touchtech AB, AVS Systeme AG and Intelliconcept AG hoped the extended software capabilities realized within the MERCO project would help them towards increasing sales in the company's main applications in the field such as exhibitions, showrooms, fairs, workshops and meetings. In addition to that, for AVS Systeme AG and Intelliconcept AG, most of the commercial expectations from the project were based on an extension of existing systems with low technical modification only. In short, the MERCO project aimed to provide software solutions facilitating remote collaboration using off-the-shelf hardware technologies with minimal instrumentation (see Figure 1.2). Since the deliverables promised were fixed and agreed upon at the beginning of the project, the execution of the project was clearly defined.

As I principally worked as one of the main researchers in MERCO during the first two years and a half of my doctoral study, the project had a significant influence on the research work included in this thesis. Apart from a system designed to support facilitated collaborative creative problem solving in remote collaboration settings delivered, the user study on the designed system not only suggested various design considerations for similar systems in the future but also unveiled different research aspects that I continued to explore in the other works included in this thesis.

1.3 Why designing for 2D interactive surfaces?

Humans have a long history of using two-dimensional (2D) surfaces of different sizes for communication. Thousands of years ago, humans already made large-scale paintings and writings on the walls of caves, tombs and temples that allow multiple people to view at a time. At smaller scales, ancient humans inscribed their writings on paper and flat surfaces of grave stones. After thousands of years, although technologies have gone through disruptive changes, 2D surfaces are still the most standard form factor for us to present information and to mediate our communication with each other. Not being limited to only physical surfaces like paper, metal or concrete, we are now interacting with computer screens, which are mainly in the form of 2D interactive surfaces, almost at anytime and anywhere in our daily life. Computer screens allow us to work in various settings. Large displays allow us to perform tasks in collaborative settings, from viewing information on public displays to collaborating for problem solving such as data analysis, group brainstorming and remote surveillance. Smartphones, tablets or laptops are suitable to perform tasks in individual settings such as writing or media editing due to their smaller screen sizes. Furthermore, those surfaces can also be combined to effectively assist people

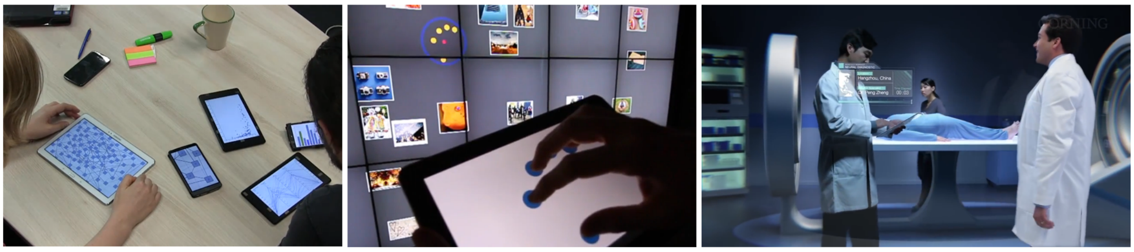

Figure 1.3: Left: Using multiple tablets for data visualization [85]; Center: Using tablet for remote interaction with a large display [152]; Right: An envisioned scenario of combining mobile devices and large display in remote collaboration in ”A Day Made of Glass”

in several tasks. For instance, tablets and smartphones can be used to interact with large displays from a distance to reduce users’ fatigue while maintaining overviews on the displays [93, 86, 112] (see 1.3). Tablets and smartphones can be combined to assist complex data analysis and sense making activities [13, 85, 97, 152, 153] (see 1.3). Two-dimensional interactive surfaces also interweave in remote collaboration. People communicate and collaborate remotely via different settings of 2D interactive surfaces. For example, it is quite common that a satellite worker uses a tablet or laptop to video-conference with his/her hub coworkers who are collocated in a meeting room, viewing his/her video feed and sharing content with him/her on a large screen. The viral "A Day of Glass" video in 2012 [6], which envisions ways of human-computer interaction in a near future, also significantly features the use of different 2D interactive glass surfaces in various activities in our future daily life. Even though the envisioned surfaces are different from our current 2D-screen devices in terms of transparency, the similar 2D screen’s form factors and sizes can make it easy to apply current interaction techniques for 2D surfaces on them. Apart from that, this vision from the industry strengthens the argument that 2D interactive surfaces will still remain as commodity interactive platforms, at least in the near future. This is an important notice for designing systems for collaboration, especially in remote settings. Also in “Distance matters”, Gary M. Olson and Judith S. Olson mentioned “technology readiness” as one of the important characteristics for a successful remote collaboration. Technology readiness refers to the alignment of technology support among remote collaboration sides. Poor alignment of technology support and the requirements for a new technology are a major inhibitor for successful collaboration [137]. For example, a video conferencing session will be disrupted if one side cannot manage to provide the video feed of their side as the other sides could not easily know what is happening there. Similarly, even though novel interactive computing platforms such as Mixed Reality (MR) glasses or holographic screens are emerging, there is a long way to go until they will be adopted as commodity devices due to obstacles in technology and social acceptability. Therefore, only focusing on supporting those futuristic platforms might lead to poor technology support alignment and reduce technology readiness for those who own 2D-screen devices, which will still remain as commodity devices in a near future.

[6] https://www.youtube.com/watch?v=jZkHpNnXLB0

I believe that the affordances of 2D interactive surfaces in comparison with recently emerging interactive platforms like VR, MR headsets or holographic displays can still warrant their commodity in the coming years despite certain limitations. One of the main challenges for users when using those emerging platforms is interacting with and manipulating the visualized digital information. Even though information can be visualized in 3D, it is still virtual. Therefore, direct interactions with the visualization such as pointing or grabbing a virtual object with hand gestures can cause physical fatigue. The reason of this is that the interactions are in-air and have no physical contact with the targets, hence interacting body parts like hands or arms receive no physical support and muscles have to maintain a certain contraction level to keep the gestures at the correct posture and position [62]. Fatigue caused by mid-air gestures has been reported in previous systems [10, 57]. Additionally, missing physical feedback like touch feeling from the target also reduces the realistic experience of the interaction [64]. Several research projects have attempted to mitigate this issue. Some work repurposed surrounding physical spaces to couple touch interaction with virtual contents [6, 17]. However, these approaches are not suitable for nomadic workers and also suffer from certain social acceptability issues. Other approaches employed haptic repurposing on handheld devices to provide more mobile solutions to coupling touch interaction with virtual contents [148, 18]. However, these repurposing approaches are only applicable in virtual reality contexts and also likely to have social acceptability issues. Interactive surfaces can mitigate this issue. While interacting with interactive surfaces, users mostly rely on touch gestures or mouses. Either way, touch interaction between users' fingers with the surfaces or the nice cause less fatigue than mid-air gestures, making it more efficient in certain tasks [8, 70]. Leveraging this advantage, recent research combined interactive surfaces with mixed reality to provide more efficient interactions with virtual information [54, 156]. This promises new interaction techniques with interactive surfaces rather than an end for this interactive platform.

Compared to 2D surfaces, those emerging platforms have more display dimensions, thus offering more realistic and immersive information visualizations. However, they may introduce issues in collaboration where participants have different perspectives on the shared workspace visualized by the platforms, thus see different content and can be occluded to what is seen by the others. This can reduce the common ground in collaboration. In contrast, even though information visualized on 2D surfaces is constrained to flat displays, it provides more consistent presentations viewed from different perspectives. This might help maintain a good level of common ground in collaboration as participants can share similar knowledge about information shown on the shared display.

Affordances of 2D interactive surfaces offer the aforementioned significant advantages but also have noticeable limitations that can hinder user interaction and negatively affect collaboration. Although touch interaction on interactive surfaces can reduce a user's physical fatigue, in certain tasks it is sometimes not sufficient to efficiently convey complex user gestures. For example, in the meeting of a design team on prototypes of a product, especially physical ones, team members sometimes need to express certain complex actions such as twisting or rotating the prototypes.

Those actions can be quickly mimicked by in-air hand gestures but may take more effort to be expressed verbally or via touch interaction on the screen. Another interaction limitation of many 2D interactive surfaces is the small screen sizes, commonly found on mobile devices like smartphone, tablet or laptop, which can hinder user interaction in remote collaboration. For example, many problems could arise when a satellite worker uses a tablet or a laptop to remotely collaborate with his hub coworkers sitting in a meeting and probably also working with a large display. First, the remote participant often needs to be able to view the entire collocated team's room space including all the participants and the large display. He might also need to be able to access and interact with the display like other people in the room. In order to maintain a high level of common ground and the remote person's sense of being present in the room, it is also important to convey to him the spatial arrangement of people and artifacts in the room. Small screen sizes of laptops or tablets make it particularly challenging to aggregate all of the aforementioned design requirements in one interface. Besides that, small screen sizes can also make it difficult to convey the attention of the remote person towards the collocated team. With a person sitting in the room, his attention to another person or a location in the room is easily perceivable by those in the room, as his attention movements (e.g. change looking direction) usually have sufficient magnitude to be noticed and differentiated. However, when the remote person views the collocated group on a tablet or laptop without appropriate presentations, the small screen of the device can make his attention incorrectly perceived by his team. This can lead to the reduced presence of the remote person in the collocated team.

Two-dimensional interactive surfaces not only offer several possibilities but also imply various challenges in designing effective interfaces for remote collaboration. Therefore, designing for this platform requires to not only leverage its strength but also mitigate the weaknesses. Furthermore, approaches aiming to mitigate the weaknesses should balance between enriching user interaction and maintaining the commodity of the platform, ensuring a high level of technology readiness in remote collaboration. For example, solutions for extending the interaction space of mobile devices should not reduce their mobility, thus should require minimum hardware instrumentation. Driven by this design mindset, works included in this thesis focus on exploiting new interaction possibilities with the given form factor of commodity devices combined with lightweight instrumentation if needed. Details of how the included works are in line with this design mindset will be further described in the paper summary section.

Chapter 2

Background and Related Work

As this thesis includes designs that support workspace awareness in remote collaboration, it is important to first provide readers brief theoretical knowledge about this type of awareness. In this part, the thesis will consequently go through three definitions: workspaces, awareness and workspace awareness. Following that, the thesis will cover how this type of awareness has been supported in different remote collaboration settings using interactive surfaces.

2.1 What is workspace awareness?

To clarify the meaning of “workspace awareness”, I believe that it is necessary to explore the meaning of “workspace” and “awareness” separately, in the context of collaboration, before discussing a shared understanding of the compound term.

2.1.1 Workspaces

Derived from a definition by Gutwin and Greenberg [50], a workspace is a bounded space where an agent (e.g. a person) performing a work can see, generate and manipulate artifacts related to the work. In collaboration, there are multiple people working on a task. The workspace where people in the collaboration can perform the aforementioned activities is called shared workspace. For example, when a team is having a brainstorming session, the shared workspace can be the table surface or the wall where the team members put up sticky notes for discussion. Likewise, together with the development of computing technology, shared workspaces can be digitally interactive. For example, nowadays a shared workspace can be a large display where multiple people can analyze data or carry out emergency response together. When it comes to remote collaboration, based on the existing definition of shared workspaces, a remote collaboration’s shared workspace is the space where involved people can remotely perform activities related to the work. For example, when two people remotely draw together, the drawing canvas shown on each person’s device (e.g. desktop, laptop, tablet or smartphone) presents the shared workspace.

2.1.2 Awareness

Despite having been investigated from early on in the history of CSCW, there is no single definition of "awareness" that works well for the larger community of HCI/CSCW. Ensley [32] defined "awareness" as knowledge created through interaction between an agent and its environment — in simple terms, "knowing what is going on". Gutwin and Greenberg [50] derived four basic characteristics of awareness from literature [1, 32, 105].

1. Awareness is knowledge about the state of a particular environment.
2. Environments change over time, so awareness must be kept up to date.
3. People maintain their awareness by interacting with the environment.
4. Awareness is usually a secondary goal - that is, the overall goal is not simply to maintain awareness but to complete some task in the environment.

In a CSCW context, Schmidt [129] mentioned that "awareness" usually does not refer to some special category of mental state existing independently of action but to an actor being or becoming aware of something or someone. One of most widely-cited definitions of "awareness" was suggested by Dourish and Bellotti [30] in which it is referred to as an actor's understanding of the activities of others, which provides a context for the activity of that actor. In addition, awareness also means that actors will adjust, align, or integrate their activities tacitly and unobtrusively to the thing or the person they are aware of [30]. For example, when multiple users are working with a shared text editor, each user's presence in the system is represented by a cursor, often having a distinct color. When a user sees the cursor of another participant at a place in the document, he/she knows that the other user is working at that place and will try to avoid manipulating there in order to avoid conflicts. We can see that the user's awareness in this example complies with the four characteristics derived by Gutwin and Greenberg. To keep being updated about the knowledge of a state of the editor (i.e. locations of users' cursors), the users need to at least look at the editing view or even click at the symbol of a user to locate his/her cursor. However, being aware of the cursors of other users is not the main goal here, rather, the main task is to finish a writing together. Studies also showed that in distributed settings, facilitating users' awareness on their remote counterparts can improve the perceived connection among each other and make them more engaged in their mutual communication [56, 101]. While there are several types of awareness in collaboration that have been investigated such as conversational awareness [21], casual awareness [11] and situation awareness [42], works included in this thesis focus on supporting workspace awareness in collaborative systems, especially in remote settings.

2.1.3 Workspace awareness

Gutwin and Greenberg [50] defined workspace awareness as up-to-the-moment understanding of another person's interaction within a shared workspace. Gutwin and Greenberg further discussed boundaries of this concept. First, workspace awareness is an understanding of people in the workspace, rather than just of the workspace itself. Second, workspace awareness is limited to events happening inside the workspace.

Third, workspace awareness is a specialized kind of situation awareness - where the situation comprises the other team members interacting with the shared workspace.

Three basic information elements that make up workspace awareness are "who, what, and where" questions about other team members and their activities [50]. When we are working with our collaborators in a shared workspace, these elements reflect essential knowledge we need to possess: who we are working with, what they are doing, and where they are working. "Who" awareness is the knowledge that there are others in the workspace, who they are and the authorship mapping between an action and the person carrying it out. "What" awareness covers the understanding of what artifacts on the workspace another person is working on or even about to interact with. "Where" awareness covering location, gaze, view and reach refers to understanding of where a person is working, where they are looking, what they can see and where in the workspace they can change things.

Workspace awareness can be gathered through several channels. The most prevalent means of communication in which workspace awareness can be formed is verbal conversation where people explicitly state what and where they are doing something or are about to do. In addition, bodily communication such as the collaborator's position, posture or movements of heads, arms, eyes, and hands provides rich sources of information about people's interaction and intention. For example, people can easily rely on the pointing direction of a person's hand to know what he/she is referring to. Or when a person turns his/her head, other people can estimate his/her next focus. Moreover, workspace awareness can also emerge from positions, orientations and movements of artifacts in the shared workspace. When people interact with the workspace, this information will keep changing, which after a while collaborators can use to derive understandings about the others' work in the space. For example, Gutwin and Greenberg provided an example of two pilots in an airplane cockpit to illustrate the role of artifacts in the shared workspace in maintaining collaborators' workspace awareness. In particular, when a pilot moves a control gear, even if the other pilot could not see such at-moment action, by observing the changes in the gear's location a while later, he/she could know about the aforementioned activity of his/her coworker. Another example in the context of a brainstorming session, when a sticky note is moved to another location, although some team's member can miss such movements due to certain distraction, he/she can still be informed about the changes on the shared workspace when seeing the change of location of the note.

Workspace awareness can benefit different aspects of collaboration. First, workspace awareness helps collaborators manage "coupling", which Savaldor et al. [127] call the degree to which people are working together. When people are sufficiently informed about the statuses of their collaborators on the shared workspace, they tend to collaborate more tightly with other people and can exploit the opportunity to share and discuss or make decisions. Second, workspace awareness helps simplify communication among collaborators. While interacting with the workspace, people usually perform deictic references, pointing or gesturing to indicate a noun used in conversations, and use many utterances like "this one", "that one", "here" and "there" [131]. This allows communication among collaborators to be much more efficient as verbal communication can be shortened. Gaze awareness, knowing where another is looking at, also helps one quickly perceive the attention or intention of another without the need of explicit expressions through verbal or bodily communication. Having good workspace awareness also helps participants better coordinate the work. Knowing or being able to predict where and what another is working on, a person can choose a task and a location in the workspace to avoid conflicts, and reduce misunderstandings [45]. These are the benefits that workspace awareness can provide to reduce effort, increase efficiency, and reduce errors in collaboration activities [50]. These benefits were not only observed and leveraged in analog teams but also in cross-device co-located collaborations to improve users' experience and efficacy in digitally sharing information among each other [94, 98].

Apart from improving team efficiency, workspace awareness has also been argued to improve "presence" - the sensation that people have whenever they feel that they can be perceived by others in whatever they are doing, including their experiences of the others [41]. Frößler argued that the perception of presence can be likely the result of the extensive exchange of awareness information, making one believe that others know what he/she is doing [41]. This perception of presence is very important to group communication as future communicative practices might be altered according to how people perceive their current communication [41]. For example, a person feeling he/she can be perceived by others, will tend to elicit more communication cues , which will foster the collaboration process. In contrast, if he/she feels not well-perceived by others, he/she might be reluctant to express himself/herself later on, thus reducing team comprehension. In short, although workspace awareness is not the primary goal in communication, supporting it is crucial for seamless and efficient collaboration.

2.2 Related work

In co-located settings, collaborators can typically take workspace awareness cues for granted from the environment as people and artifacts can be easily seen. However in remote collaboration, a collaborative system needs to principally present the information of the presence and current activities of their remote collaborators on the shared workspace via digital cues. Even though there is no systematic categorization on the representations of such cues, by surveying workspace awareness supports in the literature, they can be considered to range in a spectrum of two extremes: mediated and unmediated forms.

- Mediated forms: As defined by Kirk et al. [77], mediated forms are representations that require users' effort to extrapolate by "decoding" and aligning them with the actual information they represent. For example, when looking at a color-coded cursor on a shared text editor, a user first needs to map from its color to a corresponding user. Second, based on the state of the cursor (e.g. moving or idle for a while), the user will need to guess whether the corresponding user is actively working on the document. This concept can be found not only in shared text editors but also in creative collaboration. For example, to support collaborators' workspace awareness in information co-curation, Co-Curator [118] visualizes a pulsating dot to inform everybody about the current point of interest of the corresponding collaborator corresponding on a timeline of inspiration sources. Another example of this is by using Mediated representations are not limited to visual forms. For example, Gutwin et al. [51] used synthesized chalk sounds to inform users about activities of their collaborators on parts of the shared workspace they cannot see.

- Unmediated forms: Opposite to mediated forms, unmediated forms require almost no efforts from the users to extrapolate the information they represent. To achieve this, these forms typically employ representations that are realistically consistent with the original information. For example, overlaying hand images of a remote instructor on the workspace of a local worker would require the worker almost no effort to "decode" to understand the instructor's hand gestures [22].

In general, visual cues still play a dominant role in either representation forms in supporting workspace awareness in collaboration using 2D interactive surfaces. To represent the presence, activities and intentions of collaborators on the shared workspace, effective approaches are to visualize movements and actions of users conveyed via gestures of different body parts (e.g. hands, figures, torso) and even gaze (conveyed via head and gaze). Next, we unpack how bodily gestures and gaze have been exploited to support workspace awareness in remote collaboration and position our contributions in the field.

2.2.1 Supporting workspace awareness through bodily gesture visualization

As workspace awareness is the knowledge of what a collaborator is doing on a location in the shared workspace, one of the most intuitive approaches in supporting it in remote collaboration is to present the remote collaborator's body or gestures interacting with the workspace to their coworkers. VideoDraw [140] is one of the earliest designs for this approach. The system allows remote collaborators to see not only each other's drawing but also the accompanying hand gestures on a desktop screen. Thus, a collaborator knows not only what his/her partner is doing on the shared workspace but also where he/she is working. The authors extended this concept to large displays with VideoWhiteboard [139] where users on each side can see the shadows of their remote partners working on the shared drawing canvas shown on the display (see Figure 2.1a). Users can see not only the hand but also certain bodily gestures of their partners conveyed by the shadows. A similar approach was employed by Everitt et al. [33] in Distributed Designers' Outpost; a system supporting brainstorming using sticky notes in remote collaboration settings. Compared to VideoWhiteboard and Distributed Designers' Outpost, ClearBoard [68] further improves workspace awareness by supporting gaze communication using the metaphor of working through a glass window (see Figure 2.1b). The system overlays the shared drawing canvas on the video feed capturing the entire upper body of the remote partner working the shared workspace. This allows collaborators to perceive not only each other's hand gestures but also where on the shared workspace they are looking. HoloPort [82] applied a similar approach to present full-body gestures of collaborators on a system supporting problem solving in video communication. Leveraging the development of capturing technologies such as depth-sensing camera, 3D-Board [160] blends the 3D point cloud of the remote collaborator captured by Kinect cameras[1]. onto the shared workspace shown on a large display using the same metaphor as ClearBoard. Similarly, instead of showing 2D image of hand gestures like VideoDraw or 3D point clouds, ShadowHand [151] displays a stylized 3D reconstruction of the remote collaborator's hands on the shared workspace (see Figure 2.1c). Furthermore, the study with ShadowHand shows that either point-cloud or stylized 3D hand presentation made users feel more personally engaged with their remote coworkers than using a pointer. Instead of using the metaphor of working on a glass window, CollaBoard [84] overlays the video of the remote collaborator onto the shared workspace, mimicking the side-by-side configuration when people work together on a whiteboard in collocated settings. ImmerseBoard [61] supports both the working-on-a-glass-window and side-by-side metaphor to facilitate workspace awareness in remote collaboration. Different from CollaBoard, in the side-by-side mode ImmerseBoard affixes the video of the remote collaborator at one side of the screen and extends his/her hand from the body to the target location. CamRay [5] investigated two configurations to present the remote collaborator's video on large displays in remote collaboration. The authors found that when the video follows the position of the remote person in front of the display, it helps participants perform collaboration faster as they know what each other is doing, thus shortening verbal communication as people use more deictic references. Several works also explored different visualizations of distant people's bodily and hand gestures on the shared workspace in remote collaboration using horizontal interactive surfaces [69, 116, 125, 158, 159]. Replicating movements of physical artifacts was also explored to support workspace awareness on tangible interactive tabletops [89, 113]. As location flexibility is becoming prevalent in people's work conditions, there have been efforts spent on improving workspace awareness in collaboration using mobile devices such as laptop or tablet, which typically possess small screens. The primary collaboration scenario of this topic research is a worker using a mobile device to communicate and receive instruction from a remote expert. The worker's workspace is captured and displayed on the devices of both the worker and the expert. Fussel et al. [39] explored a setup where the expert can elicit gestures mediated by drawing on the video feed which can be seen by the remote worker while giving the instruction. They showed that such annotation can improve workspace awareness in such collaboration as the worker perceives the target indicated by the expert and how to manipulate it more quickly. This approach was also adopted in other remote collaboration systems supporting mobile workers [29, 40, 76] (see Figure 2.1d). SEMarbeta [16], besides using sketch-based annotation, also employed unmediated hand gestures by capturing and displaying images of the instructor's hand on the

[1] `https://developer.microsoft.com/en-us/windows/kinect`

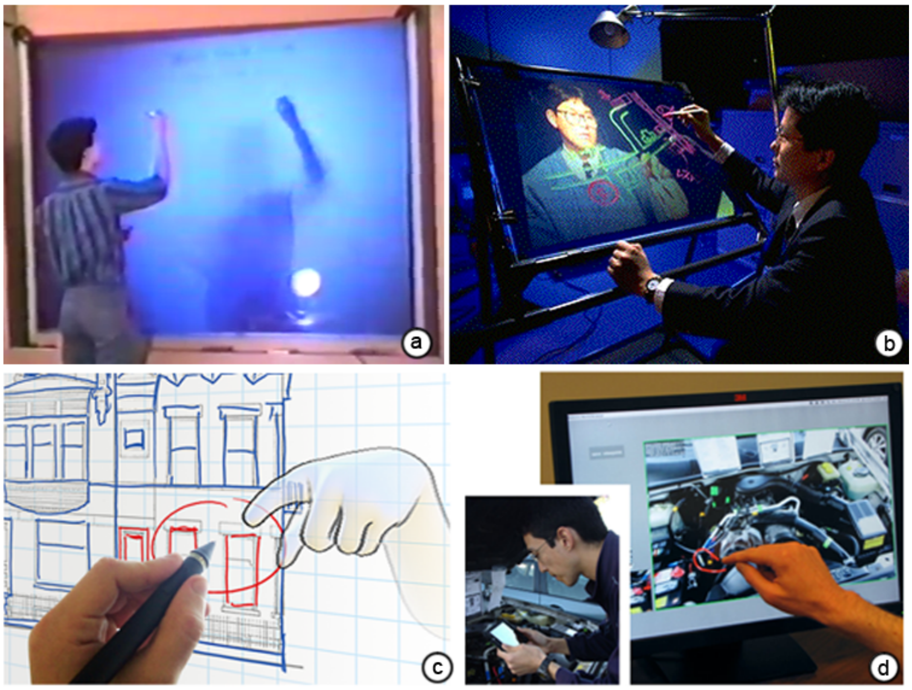

Figure 2.1: (a)VideoWhiteBoard: showing shadows on the remote collaborator (b)ClearBoard: overlay images of the remote collaborator on the shared workspace; (c) ShadowHands: blending stylized visualization of hands with shared workspace; (d) System of Gauglitz et al. [40] allowing mobile user viewing hand gestures on mobile devices represented by sketch-based annotation on shared workspace created by a remote person using a touch screen

remote worker's device. However, due to spatial discrepancies in the hardware setup on the instructor's side, benefits of showing unmediated hands of the instructor were hard to observe. Also leveraging showing images of the remote collaborator's hands on the video feed, BeThere [135] is a proof of concept that employs several external sensors to extend the interaction space of a smartphone beyond the device's small screen, allowing users to perform hand gestures around the device. However, there still remain unsupported scenarios in existing works. One common scenario is a traveling person remotely collaborating with a team sitting in a room and working in front of a large display, which is referred to as hub-satellite settings in the literature [81, 143]. Designing interfaces on mobile devices for the remote person in such cases to view the team's space, work with the shared workspace on the display, and maintain awareness on other team members still remains unsolved. Additionally, in expert-worker remote collaboration similar to the contexts investigated by Fussel et al. [39] and SEMarbeta [16], previous approaches had to rely on stationary or heavily instrumented hardware to capture images of the expert's hands, reducing the commodity and mobility of the system. Works included in this thesis aim to address those issues, providing nomadic workers with lightweight solutions supporting workspace awareness in their remote collaboration practices.

2.2.2 Facilitating gaze communication to improve workspace awareness on interactive surfaces

Although gaze is often considered as a subtle non-verbal cue, it is an important referential indicator in communication [31]. Gaze, conveyed by directions of head or eyes, serves as a signal through which a person relates their basic orientation and even intention toward another [73]. Thus, supporting gaze awareness can make communication among collaborators more efficient. In approaches capturing and overlaying images of the remote collaborators on the shared workspace such as ClearBoard [68], HoloPort [82], 3D-Board[160] and ImmerseBoard [61], gaze can be inherently achieved due to the alignment of the collaborators' images and their locations on the workspace. This approach was also extended for more complex setups to convey gaze in remote collaboration environments using multiple interactive surfaces like MultiView [103], MMSpace [109] or even tabletop collaboration [59]. However, on personal computers such as desktop, laptop or tablet, this approach might not work due to small screen sizes. Rather, designs on these platforms focus on visualizing gaze information provided by an eye tracker data as graphical stimuli on the shared workspace to attract the attention of the remote person [2, 27, 26]. This approach was also employed for tabletop remote collaboration to facilitate gaze awareness on physical artifacts involved in the collaborative tasks [3]. Also supporting interpreting gaze on physical artifacts in the shared workspace, ThirdEye [110] is an add-on display attached on tablet screens that conveys the remote worker's gaze into the 3D physical space. It projects a 2D graphic element, controlled by eye tracking data of the remote worker, onto a hemispherical surface that looks like an eye. As we can see, those above works focused on conveying gaze on only one type of subject at a time (e.g. remote partner or artifacts on a shared workspace). This leaves settings, such as a satellite worker remotely having a meeting with his/her hub coworkers sitting in a room discussing physical prototypes on a table, unsupported. This thesis also aims to fill this gap.

2.2.3 Summary of related work

Even though there have been plenty of efforts trying to improve workspace awareness in remote collaboration by leveraging various mediated and unmediated cues, there are still considerable distributed collaboration settings that remain under-explored. First, there is still a lack of exploration into how workspace awareness support affects collaboration dynamics of asymmetric teams such as hub-satellite settings, having different roles and using hybrid infrastructures ranging from large displays to mobile devices. As such settings are becoming common, this thesis provides investigations to pave the way to designing workspace awareness supports according to collaborators' roles and the devices they use. Second, nomadic workers using commodity personal computers like laptop or tablet are still receiving inadequate support to richly express their non-verbal communication such as bodily gestures or gaze in collaboration with their remote collaborators. This thesis thus aims to fill this gap by providing new solutions that require minimal hardware instrumentation to help them better convey such cues, in order to improve workspace awareness in their collaborative work.

Chapter 3

Research Methods

All papers included in this thesis are design-oriented works, which typically involved the understandings of user behaviors and problems they encountered and iteratively designing and critiquing design solutions to solve the defined problems. A user-centered design process was employed in the works included in this thesis. The Interaction Design Foundation defines user-centered design as interactive design process in which designers focus on the users and their needs in each phase of the design process in order to create highly usable and accessible products for them [35]. Within my research, user-centered design was particularly employed as follows. First, understandings about workspace awareness and how it is an issue in existing solutions were acquired through literature review, field study, focus group observations and interviews. Identified problems were then discussed and distilled to narrow down to the most worthwhile ones to tackle. After that, brainstorming sessions were carried out to find the most appropriate design solutions to solve the problems. The designs then went through iterations of prototyping and critiquing. The final prototypes were formally or informally evaluated. The outcomes of this process are not only novel solutions to mitigate existing problems in remote collaboration, but also insights about why they work and potential issues from HCI perspectives. These served as design considerations shared with the community to design future systems and to further study relevant aspects.

3.1 Understanding the theoretical background of collaboration

Literature review was the main tool to acquire fundamental knowledge about factors, properties and issues of collaboration, especially in remote settings due to a quite established research body in this field. During this process, the theoretical background related to workspace awareness was studied and reflected in existing remote collaboration systems. This process also helped identify which areas in the topic were still uncovered and open for later research in this field. The literature review started with the classic "Distance matters" [108] to overview the most common theoretical aspects in remote collaboration and then gradually extended to publications referencing to this work or cited by it.

3.2 Eliciting design requirements

Identifying concrete requirements is crucial for a design process. Typically, design requirements depend on particular contexts of users, such as what tasks they are working on, which devices they are using, who they are working with or what are the conditions of their environments. To adequately gather such contextual information, various techniques were employed in this thesis.

3.2.1 Literature review

Several typical issues in user experience and interaction related to workspace awareness in remote collaboration were systematically documented in previous works such as [108, 50, 30]. These provided a reliable source of references on the way to define problems for research. The issues described in this source were then reflected to existing technical solutions to see how they were supported or still remained challenging. This process was further examined regarding the introduction of the latest hardware devices, which led to new practices in remote collaboration that might not be effectively supported by the existing solutions. Finally, the entire process elicited issues in workspace awareness with regard to new computing devices and new remote collaboration practices, which were selectively considered as research problems in the thesis. More specifically, the main research topic chosen to be tackled in this thesis was how to design to support workspace awareness in remote collaboration where traveling individuals use commodity mobile devices like tablet or laptop.

When I started my research, I began my literature review on technical systems supporting workspace awareness in collaboration, especially in creative problem solving due to its relevance to the MERCO project, such as VideoDraw [140], ClearBoard [68], Designers' Outpost [80] and its distributed version [33] and CollaBoard [83]. The literature review process then extended to other articles referencing to those systems. To have an overview on all interfaces and interaction techniques for workspace awareness support in remote collaboration, I also considered other types of collaborative tasks other than creative problem solving, such as expert-worker remote collaboration on physical tasks. Likewise, I also surveyed distributed team settings typically supported by the systems in the literature. All of these provided me early ideas on which team settings as well as which hardware infrastructures are still inadequately considered when it comes to support workspace awareness. Likewise, when I had a new idea for a paper, doing literature review helped me acquire necessary knowledge about whether the problem I wanted to solve was already addressed and how my solution differed from existing approaches.

Certainly, this literature review could not conceive all lower-level research problems in this thesis at the beginning. They were actually consequentially elicited: insights on a design addressing previous problems aspired the following ones. Other techniques such as user observations and interviews were also combined with literature review to elicit worthy research problems from users' perspectives.

3.2.2 User observations

Observing users is usually considered as an effective way for researchers or designers to build their user empathy, helping them to think from the users' perspective [154, 79]. Sometimes, it can be difficult for humans to explain what they do or do not do, or even describe accurately how they achieve a task. User observations thus complement other techniques such as interview, focus group, or literature review in eliciting user requirements [119]. Works in this thesis were based on user observations in different stages of a design process. At the beginning of my study, user observations were carried out in a field study in remote meetings of design teams in companies to identify certain collaboration practices that were not documented in the literature. Besides that, user observations were also carried out in formative and summative evaluations [130] conducted using the final prototypes. This kind of observations can reveal users' behaviors and interactions with the evaluated systems or with their remote partners through the systems, that even the users did not notice, informing design considerations for future improvements. In this thesis, a work is typically devised from observations in the formative or summative evaluation of a previous one.

3.2.3 Interview

Even though there has been the perception that users do not know what they want [99], interviews are essential sources of requirements in user-centered design [149] as users know most about the problems that they are facing [55]. Interviewing is a great way to gather design requirements from the users when many of them cannot be found in the literature or observed in field studies, especially with regard to new technologies. As computer technologies used in the literature are usually out-of-date compared to the ones being used in real life, interviewing can elicit new requirements from the perspectives of users using such devices. Besides that, interviewing participants in prototype evaluations usually provides researchers or designers opportunities to listen to the users about how the prototypes or related issues can be further improved. Similar to user observation, such interviews helped inspire me for new ideas after each project.

Interviews can be done in various settings, including structured, semi-structured and unstructured [72]. While structured interviews strictly follow a planned procedure using the same wording and sequence for all participants, unstructured ones, on the other side, are flexible not guided by any pre-defined plan and can be different among participants. In my research, I mostly applied unstructured interviews which typically had some main questions to ask all participants but more details can also be added depending on the direction of each interview. This type of interview allows me to identify user problems as well as explore usability and user experience adequately. For user need elicitation, after the literature review and user observations, I typically had rough ideas about what problems the users might encounter and thus could define a set of basic questions to guide the interview process. Then, in each interview, depending on the answers of each participant for those questions, I investigated further in spin-off details using additional questions. Similarly, when it came to the evaluation of a system (see the Evaluation section below), as a researcher, I typically had a set of most crucial questions to explore usability and user experience aspects of a designed interface. Then, depending on what I observed on the participant's interaction and behaviors during the user study or what the participant answered for the main questions, I would pose different additional questions to better understand the impact of the device on users.

3.3 Design and Prototyping

Prototyping is the stage where the tangible realizations of a design are created, which is a critical step in a design project. Prototypes serve as a medium for designers to reflect on their design, explore different alternatives, demonstrate the ideas with clients and users, and investigate findings related to user interactions and experience. Prototypes have different resolutions, typically ranging between two extremes: low-fidelity and high-fidelity prototypes. Low-fidelity prototypes, usually in the form of sketches or paper presentations, capture early conceptual ideas. These types of prototypes are often easily and quickly made, allowing designers to explore different design possibilities and identify early problems in design alternatives. A variance of sketch prototypes that was also frequently used in the design process of the works included in this thesis is storyboarding. Storyboarding is a technique where designers sketch how a user might progress through a task using the device being developed, or a series of scenes showing how a user can perform a task using the device [119]. High-fidelity prototypes are often visually and functionally close to the final product. This type of prototype is usually used for more mature testing and user experience evaluation [119]. More profound interaction and design considerations are often conceived by evaluating high-fidelity prototypes. In this thesis, high-fidelity prototyping includes crafting hardware and developing software.

3.4 Evaluation

Evaluation is essential to verify if a design solution could solve the provided problem. Several techniques were employed to quantitatively and qualitatively evaluate systems included in this thesis. In general, the user studies with the systems in the thesis were conducted in laboratory environments due to constraints related to infrastructure and human resource. I adopted and customized task designs of established related works for my studies with similarities in collaboration settings that the works aimed to support. Likewise, established and suitable quantitative measures from related literature were also adopted and customized. Basically, task completion time, accuracy and perceived workload when users collaborate to finish a task were used to measure the performance of the system. To measure perceived workload, NASA Task Load Index questionnaire [58] and time to respond were employed due to its common use in literature. System Usability Scale (SUS) [12] and customized questionnaires were used to quantitatively measure the usability of the developed system and user experience. Apart from that, qualitative feedback from the users was also collected through post-study interviews and open questions in questionnaires to gain deeper understandings about user interactions and experience with the system, and why the system works or does not work. I also employed an informal heuristic evaluation [104] for the works in this thesis. Occasionally, developed systems were demonstrated to experienced researchers in the field visiting our lab. After watching and trying out the systems, the researchers provided critiques on potential usability and user experience problems in the designs. This was very helpful because besides fixing issues in the design, it also inspired new research ideas.

3.5 Ideation

Several meetings and brainstorming sessions among research project team members for ideation were held in the different phases of the research process of the work, from distilling user requirements, defining research problems, finding design solutions through to evaluation. Such sessions are important for creativity activities like design and research because they allow the team to gather different perspectives of team members. When team members exchange their ideas in such sessions, the team will gain greater accumulated knowledge, helping them to more likely arrive at a better solution than one produced by an individual [147]. In most cases, ideation meetings were conducted in facilitated manners where I, as the project leader, often acted as a facilitator, making sure that the team always stayed on track with the expected outcome given at the beginning of the session. This facilitation is especially important as most of the brainstorming sessions were conducted in distributed settings due to the cross-institution collaboration nature of my research.

Chapter 4

Summary of Research

In this chapter, I first provide a summary on the thesis' contributions and my research approach to clarify the connection of the pieces of work included in it. Next, I provide more details about how each piece of work complements another and how it contributes to improve workspace awareness in remote collaboration using 2D interactive surfaces. Finally, summaries of included papers are provided to offer readers an overview on the research process and the main outcomes of each paper.

4.1 Contribution overview and research approach

The contributions of this thesis are a set of working prototypes that solve different issues in remote collaboration. Furthermore, reflections on the designs of the prototypes regarding user interaction and communication using the proposed systems serve as considerations for designing future collaborative systems. The research projects were mainly carried out in sequence, where user feedback from a previous project

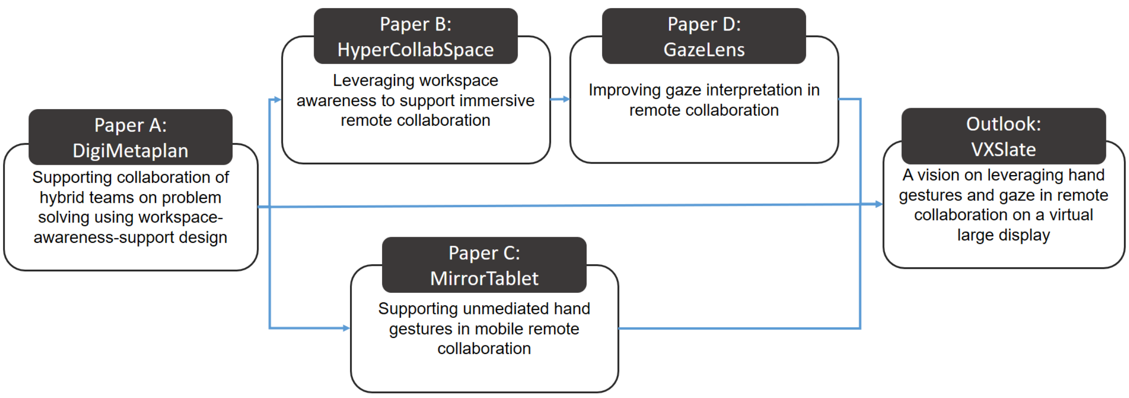

Figure 4.1: Summary of the concepts included in this thesis as well as the consequential relationship among them. With each arrow, the ending concept was inspired by the outcomes of the starting one. The "Outlook:VXSlate" is a synergistic visionary concept and presented in the "Outlook: Virtually Extendable Slate" section

sheds light for the following one, which in turn aims to address user interaction problems encountered in the former. The contributions of this thesis focus on two main directions:

- Exploring the effects of workspace awareness cues in novel remote collaboration settings
- Designing interfaces and interaction techniques to improve workspace awareness in remote collaboration.

My research approach was first to explore how workspace awareness affects collaboration dynamics of a distributed team in creative problem solving, as this type of collaborative task typically involves rich multi-channel communication and interaction: among team members and between team members and artifacts, potentially revealing interesting underexplored research problems. Besides that, this is also a common collaborative activity in industry, making my research more pragmatically beneficial for the real workplace. I also chose to study such collaborative tasks in hub-satellite teams in a structured process, as these settings can be easily found in real-world practices but still remain underexplored. In this study, I only focus on designing collaborative interfaces on commodity touch-screen devices such as smartphone, tablet and interactive whiteboard. Thus, I mainly utilized mediated cues to stimulate workspace awareness in the design caused by the constraints of those devices. Based on the insights from this study, I focus on remaining issues in workspace awareness not effectively supported on personal devices. Therefore, the later focus of this research body is how to support remote individuals to better maintain and express their workspace awareness while collaborating using mobile devices, charted in three aspects as follows:

On one side, I designed a novel concept that aims to improve remote participants' awareness on the hub's workspace when collaborating with the hub using mobile devices like tablets. The designed concept provides a spatially consistent mapping of the physical hub's space onto the device's screen, combined with mediated representations of hub's coworkers to offer the remote participant flexible interactions while adequately maintaining their awareness on activities related to the shared workspace. On the other side, I also improved the coworkers' awareness on the remote participant's activities on the shared workspace. For such purpose, I focused on better conveying non-verbal communication cues of the remote participant, elicited via hand gestures and gaze as they are effective means for humans to express their attention. Unmediated representations were leveraged to provide unobtrusive and expressive solutions. For hand gestures, I designed a compact system for capturing and presented unmediated hand gestures on mobile devices and explored the effect of the resulted unmediated representation in mobile remote collaboration. However, although using hand gestures to convey attention is a natural behavior of humans, in conversations or collaborations, people frequently do not convey their intention via hand gestures due to certain reasons (e.g. conversational pace or social acceptability). In that case, gaze is the main non-verbal cue for expressing attentional information of a person. Therefore, to round up this thesis, I designed to better support awareness of gaze direction, a more effortless and unobtrusive way for conveying attention in collaboration. I designed and evaluated an interface solution that does not require any special hardware on the remote participant's side but can still help hub coworkers effectively interpret his/her unmediated looking gaze in his/her video feed. I believe that these explorations on multiple facets can improve two-way workspace awareness between a co-located team and remote participants in partially distributed teams. In the following, the contributions of the pieces of work in this thesis are summarized. Also, Figure 4.1 briefly summarizes these pieces of work as well as their consequential relationship.

4.2 Summary of contributions

This section explains in more detail how the work included in this thesis contributes to improve workspace awareness in remote collaboration. The order of the contributions is arranged following their chronological properties and reflects the shift in my research focus: from leveraging workspace awareness in multi-device systems to supporting nomadic workers using personal computers (laptop, tablet), which I found interestingly undersupported. This direction is inline with what I envisioned in my licentiate defense about the latter half of my doctoral study.

4.2.1 Supporting collaboration of hybrid teams on problem solving using workspace-awareness-support design

Creative problem solving activities such as brainstorming are a type of group work commonly found in various organizations from universities to companies. Due to its collaborative nature, this type of activity is usually carried out in collocated settings. Collaborative creative problem solving like brainstorming has been shown to benefit from facilitated and structured processes where the groupwork is moderated by a facilitator to follow sequential phases of collaboration [138]. More specifically, the facilitator and the structure of phases can help stimulate interactions among team members to exchange knowledge, guide the process, and eliminate unnecessary time-consuming details [53]. Therefore, they keep the team staying on track, improving team efficiency [96, 107]. However, such facilitated group work is currently still limited to collocated teams using paper-based settings. In contrast, due to current traveling practices, it is quite often that some satellite workers might be away from their hub, hence they have to remotely collaborate with their hub coworkers. Although role separation and phase shifting have been observed to emerge in cross-device problem-solving collaboration [95, 102], digital tools in general still tend to neglect supports for these practices, thus not suitable to support facilitated collaborative creative problem solving in such hub-satellite settings.

To address this, I designed DigiMetaplan, a digital tool supporting facilitated brainstorming in both collocated and distributed team settings (see Figure 4.2). I derived design requirements for the tool from the literature and the Metaplan method, an established facilitated brainstorming method widely used in companies [53,

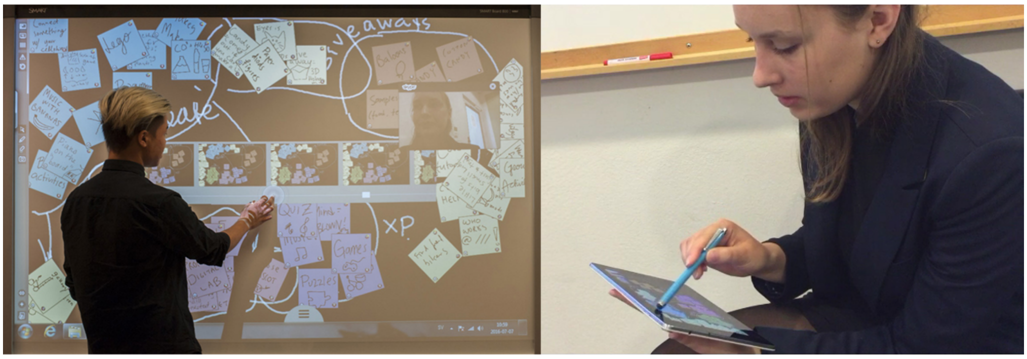

Figure 4.2: Design of DigiMetaplan supporting collaboration in both collocated and distributed settings. Left: system interface on an interactive whiteboard for a co-located team. Right: an individual remotely working with the co-located team using a system interface on her tablet

74]. The system leverages the prevalence of large interactive displays as a shared workspace (showing the shared brainstorming canvas) for hub coworkers in a meeting room and of mobile devices like a tablet commonly used by satellite workers. The design of the DigiMetaplan system is meant to support users to effectively perform facilitated brainstorming using the chosen device form factors. DigiMetaplan can effectively support role separation and phase shifting in facilitated brainstorming by adaptively leveraging workspace awareness according to user roles in each phase of the process. When team members need to focus on individual work, the system provides minimal cues to attract users' awareness to the shared brainstorming canvas. By contrast, in collaborative phases, requiring global participation and thinking of all participants, the system strategically employs various graphical tools in form of mediated presentations such as pointing highlighters, note-color changing, note manipulation, sketching and a timeline of the brainstorming history to foster team awareness on the shared workspace in order to establish a high level of common ground in the team. From the study, I could figure out how different types of digital workspace awareness cues can aid facilitators in guiding the team process. Besides effectively supporting the structure and role separation as in the conventional method, the developed digital cues also led to emergent behaviors and interaction techniques of different team members that help teams collaborate more efficiently. Furthermore, I also explored how different means to interact with the shared workspace (here, the brainstorming canvas) of various roles of users can influence a team's workspace awareness and group coordination. The contributions of this piece of work are, thus, not only the design and implementation of DigiMetaplan, a system supporting facilitated brainstorming, but also insights from a study with the system on how the designed features influenced the collaboration dynamics of distributed teams.

Even though this paper is one of the latest published pieces of work, this is my very first exploration on computer-supported collaborative system, eliciting different research problems for later papers. Besides that, this is also the only paper from my licentiate included in this thesis. Thus, it acts as a link, showing the development of my research: from exploring various types of awareness for different collaboration settings in the first half of my doctoral study to more focused approaches on improving workspace awareness for remote collaboration with involvement of satellite workers in the latter half of my study. The corresponding publication of this work is ***Paper A*** in the ***Summary of Articles*** section.

4.2.2 Leveraging workspace awareness to support immersive remote collaboration

One of the major issues that a satellite worker had in DigiMetaplan was that they could only see the shared workspace and hub coworkers working on it, but not other hub coworkers who are not nearby the board. This hindered the communication between them as the satellite worker could not see what those hub coworkers were doing, or had difficulties to differentiate who was speaking. This suggested that it might be necessary to integrate a panoramic video of the hub's space onto the interface on the satellite's side. The panoramic video can be streamed from a 360° camera, placed at the center of the hub's environment. However, integrating the panoramic video with the shared workspace on a 2D screen can be tricky. The integration needs to allow the satellite worker to freely interact with the shared workspace and still adequately maintain their awareness on the activities of the hub's people, especially in relevance to the shared canvas. Existing approaches for such an integration typically display the shared workspace and a 2D panoramic video of the hub coworkers in two separate windows [25, 126]. This approach is suitable for commodity devices like tablet or laptop. However, this presentation approach can lead to spatial disparity between the shared workspace and hub coworkers who are not nearby. For example, when one of those coworkers performs some deictic gestures by pointing his/her hand towards the shared workspace, due to the spatial separation between the shared workspace and the panoramic video, it might be difficult for the satellite worker to determine where the hub coworker is aiming at. Besides that, unwrapping the 3D space of the hub's environment into a 2D panoramic video can make it confusing for the satellite worker to interpret hub coworkers' gaze. For example, if a hub coworker sitting at a hub table looks straight forward in the panoramic video, the satellite worker might perceive that he/she is looking at him while in fact he/she is looking at another hub coworker sitting on the opposite side of the table. Another example is when a hub coworker situated afar from the shared workspace points at it while discussing some artifact, such spatial separation might lead to difficulties for the satellite worker to interpret the pointing direction. Thus, such placements of the hub's people and the shared workspace can cause overhead mental efforts for the satellite coworker to align activities of hub's coworkers with their intended location on the shared canvas. Additionally, on devices with small screen size, the images of hub coworkers in the panoramic video will be too small, and hard for the satellite worker to see. Besides that, possibilities in dealing with occlusions on the shared canvas are also limited with this interface. With hub coworkers, if their view to the shared workspace is occluded, they can move their head or even change their position to see the occluded content, whereas the satellite worker cannot do that on the 2D video of the shared canvas shown on his/her screen. All of these issues can lead to the reduction of immersiveness of the collaborative system (i.e. the satellite worker does not feel that he/she is a part of the team).

Keeping the above issues in mind, I designed a novel interface on the satellite worker's screen facilitating workspace awareness to support the immersive experience of the satellite worker. This concept will be later referred to as HyperCollabSpace and corresponds to ***Paper B*** in the ***Summary of Articles*** section. The design of HyperCollabSpace incorporates a cubic virtual environment representing the hub space. Rather than being displayed flat, the panoramic video of the hub is split into four equal segments and overlaid on four walls of the environment. A digital copy of the shared workspace is overlaid at its corresponding position on the panoramic video. This presentation helps maintain the spatial consistency of the virtual environment with the actual hub's space. HyperCollabSpace is designed to support different hub-satellite collaboration, not only facilitated processes as in DigiMetaplan. On the hub's side, HyperCollabSpace aims to support even multiple hub coworkers working with the shared workspace, instead of only one facilitator. HyperCollabSpace represents hub coworkers working with the shared workspace by 3D kinetic avatars at their corresponding locations in the virtual environment. The appearance of these avatars can be synthesized from the outlook features of their corresponding hub coworkers such as height, hair and clothing colors captured by depth-sensing cameras placed at a side of the shared canvas. Also, these avatars move according to their corresponding hub coworkers, which can be tracked by the depth-sensing cameras. This presentation allows the satellite worker to keep track of activities of hub coworkers on the shared workspace. Furthermore, it allows the satellite worker to navigate around the virtual environment in order to change his/her perspective, especially when the avatars occlude the shared board. I chose to use synthesized kinetic 3D avatars because behavioral realism was shown to have a stronger effect than appearance realism to anthropomorphic perception of users [133]. An additional reason is that using synthesized avatars will reduce the amount of data to be transmitted as the system just needs to send and receive the data of the skeleton joints' positions rather than the entire point-clouds, warranting the system to work in real-time. To support the satellite worker to observe what is happening on the shared workspace while maintaining his/her awareness on other hub coworkers, I adopted the write-on-glass interface of ClearBoard [68]. Using this view, the satellite worker will have the impression that he/she is standing on one side of a glass shared canvas, thus seeing the content on the shared workspace and the hub coworkers working with it as well as other hub coworkers on the walls of the virtual room. To ensure the content on the canvas to be readable to the satellite worker, we keep the original orientation of the contents that are considered textual and mirror the sketches. For example, the content of a note will be kept unmirrored. Even though there could be the cases where sketch-based writing can be mis-mirrored, I believe that this approach can help the satellite worker read the shared canvas' content adequately. In general, all of these interface features are designed to be incorporated on a limited screen size and leveraging touch interaction for navigation, thus suitable for commodity devices like tablet or laptop. A satellite worker can interact with the shared canvas on his/her device through pointing on the screen using touch interaction, which is represented as a color-coded pointing highlighter on the canvas.

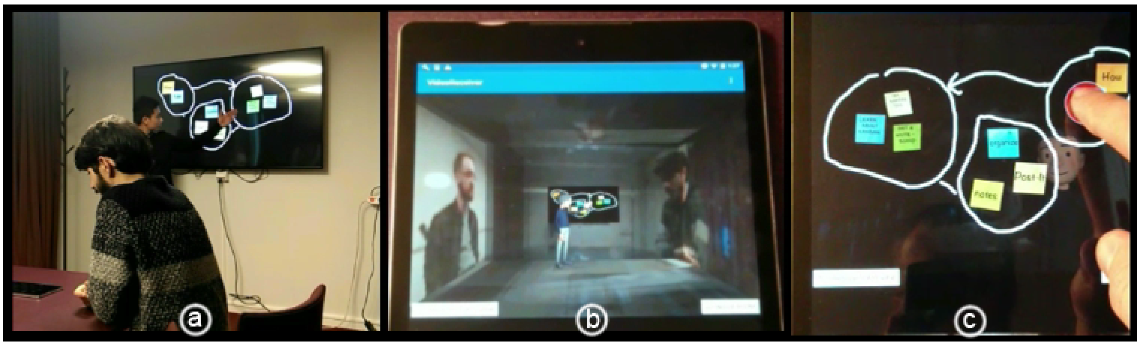

Figure 4.3: HyperCollabSpace setup. (a) the hub's space with team members sitting around a table and a facilitator working with the shared interactive display. (b)HyperCollabSpace virtual environment shown on the remote participant's tablet, panoramic video of hub's space overlaid on the four walls, the digital copy of the shared display placed at its corresponding location and the facilitator replaced by his kinetic 3D avatar. (c) In the write-on-glass mode, the satellite worker can interact with the shared display (here: pointing) while maintaining his awareness of the facilitator and other people in the room behind the transparent screen (here: the virtual walls textured by the panoramic video are not clearly seen due to the darkness of the photo caused by the light condition when it was taken)

Although the illustration in the paper consists of only one satellite worker, the design of HyperCollabSpace can support multiple individuals of this type of collaborator. Thus, HyperCollabSpace helps hub coworkers and other satellite workers (if any) maintain their awareness on where a satellite worker is working.

Due to the lack of development manpower, I could not build a robust prototype for further evaluating the proposed design. Nevertheless, I believe that the system, if fully implemented, can be explored in the following aspects to investigate the effects of the design on collaboration dynamics of a distributed team.

- **Verbal communication of team members:** Here I can employ a similar approach of measuring talking turns of participants as in DigiMetaplan as an objective quantitative indicator for the involvements of team members in collaboration. This approach is adopted from a study of Biehl et al. [9] in a similar distributed team setting, consisting of a satellite worker collaborating with a hub. Compared to conventional video conferencing or video conferencing using panoramic video feeds [126, 25], HyperCollabSpace provides more flexible navigation in the virtual space for a remote user to explicitly specify his/her current focus on certain hub members. I thus hypothesize that the interpersonal communication between the satellite worker and hub coworkers using HyperCollabSpace might be higher than in conventional video conferencing. This should be reflected in higher turn rates of the satellite worker and higher turn overlapping rate between the satellite worker and hub coworkers, as the satellite worker might be more willing to take his/her turn and even interrupt a hub coworker. However, as mentioned previously, I still have not solved the problem of presenting the satellite worker's attention on the hub's side. Therefore, the sense of presence of the satellite worker on the hub's side could be hindered, thus the satellite worker's verbal communication would not be comparable to hub coworkers'.

- **Cues of attention:** When people are paying attention to a certain subject, they normally focus their looking direction (i.e: gaze) toward it [44, 37, 71]. In face-to-face collaboration, this can be easily perceived through head pose and eye gaze. However, due to the small screen size on mobile devices, this can be hard to perceive or even tracked by eye-trackers, if the satellite worker uses conventional video conferencing. Nevertheless, with HyperCollabSpace, this can be explicitly elicited when the satellite worker navigates around to change their perspective in the virtual environment. It could be thus interesting to investigate if changes of the satellite's attention conveyed through such navigation are comparable to changes of attention of hub coworkers conveyed via their gaze. This will provide insights into whether the designed navigation in virtual environments would be an effective tool for the satellite worker to maintain his/her awareness on the hub coworkers and the shared workspace.

- **Immersion:** Immersion of the satellite worker is his/her feeling of being in the hub's space. This information can be gathered through his/her subjective feedback in questionnaires. For example, a satellite worker can indicate his/her level of agreement on Likert-scale choices [92] for statements such as "I felt like I was in the room of the hub". Immersion can also be indirectly reflected via how the satellite worker perceives the spatial arrangement of people and artifacts in the hub space represented on their device. Interfaces that allow users to quickly and easily understand relative locations between artifacts and people in the hub side as well as the directions of their activities should provide a better immersion. This can be measured via questionnaires (e.g. Likert-scale based agreements on statements like "I could easily understand whom/what a hub coworker is referring to") or in comparative tasks (e.g. setting up highly controlled tasks where the satellite worker needs to specify targets a hub coworker is referring to via gaze or pointing gestures). I hypothesize that HyperCollabSpace will outperform conventional video conferencing interfaces in this aspect due to integrating the panoramic video and the shared workspace in a spatial consistence with the hub's space. Besides that, the write-on-glass view can allow a remote participant to be better aware of activities of hub coworkers not working with the shared workspace while following what is happening on it. Even if none of the hub coworkers could have this ability, I hypothesize that this might provide the satellite worker better immersion compared to conventional interfaces separating the panoramic video and the shared workspace [126].

- **Users' attention in different collaborative tasks:** HyperCollabSpace provides an environment that is suitable for different collaboration settings: around-table team meeting, facilitated group work as in DigiMetaplan, non-facilitated group work (e.g. all team members working with the shared whiteboard in brainstorming or collaborative sketching), or group presentation (e.g. a person presenting to other hub people, using the shared workspace for showing the presentation slides). I believe that participants have different patterns of attention in different collaboration settings. The HyperCollabSpace interface, thus, could allow us to explore remote satellites' attentions in those settings, serving as design considerations for future collaborative systems.

Even though not formally evaluated, the design was showcased to senior researchers in the field of virtual reality visiting our lab. Feedback and critiques I received from these sources can be considered as informal heuristic evaluation. One significant observation was that the hub coworkers can only perceive the attention of a satellite worker when working with the shared workspace (through the color-coded pointing highlighter). However, when the satellite worker is paying attention on a hub coworker, the current design does not support conveying such information to the hub side. This led to a work that will be described later, aiming at conveying the attention of the satellite workers on different people and objects on the hub side.

4.2.3 Supporting unmediated hand gestures in mobile remote collaboration

From our observations and user feedback from the study with DigiMetaplan, one of the major disadvantages of the satellite worker that caused his/her reduced presence was potentially the limited experessibility of his/her video feed shown on the hub side. In fact, due to the limited field of view of the built-in camera of the tablet, hand gestures of the satellite worker could not be captured and presented adequately in the video feed. Even though touch interaction could convey certain deictic gestures, several hand expressions such as in-air gestures above the screen or gestures mimicking object manipulations cannot be easily expressed via commodity tablets. Suggested by previous work [78], visualizing unmediated hand representations on the satellite worker's tablet could be a solution for richer expressibility of the satellite worker. However, form factors of commodity mobile devices do not allow to effectively capture unmediated hand gestures while interacting on and above the device screen. This led me to two research questions. First, how to design solutions to facilitate unmediated hand gestures in mobile remote collaboration, without sacrifying the platform's mobility? Second, would unmediated hand representations affect remote collaboration using mobile devices? I attempted to answer these two research questions in one of the works included in this thesis, named MirrorTablet. Informed by the literature, I chose to explore the research questions, especially the second one, in the context of remote collaboration on physical tasks (e.g. a technical expert remotely instructing a worker in machine assembly or repairing), as this type of collaborative tasks typically requires rich non-verbal cues to supplement verbal communication, and it is thus more likely easier to analyze the effect of hand gestures.

MirrorTablet offers three contributions. First, the piece of work provided the design and a working prototype of a system, including hardware and software components, that captures and presents unmediated hand gestures on and above a tablet screen in remote collaboration. Leveraging the form factor of commodity

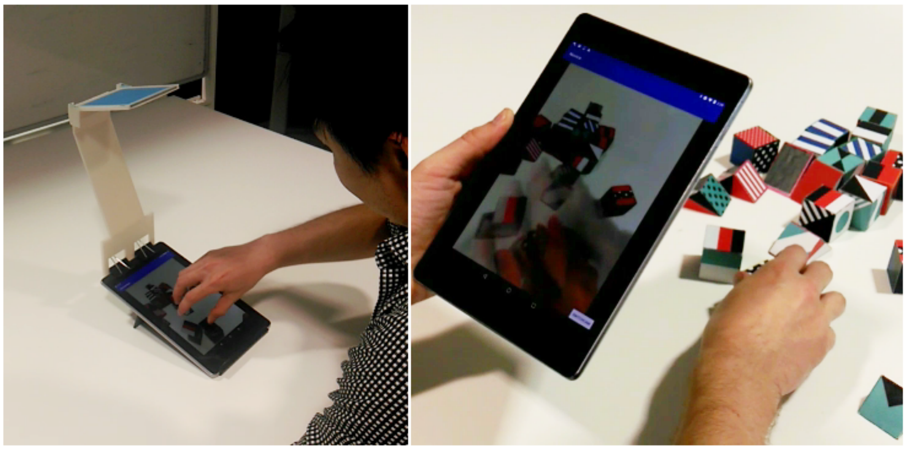

Figure 4.4: MirrorTablet system. Left: proposed hardware prototype that is lightweight and requires minimal instrumentation on commodity mobile devices, in fact, this can be designed to be foldable as a tablet cover. Right: Remote participant viewing unmediated hand gestures captured by the proposed prototype.

mobile devices like tablet or smartphone, MirrorTablet introduces a foldable structure holding a mirror to reflect users' hand gestures on and above the device's screen, which is then captured by the built-in front-facing camera of the device. Captured images are then processed to extract unmediated hand gestures, which will then be overlaid on the video feed shown on the device of the remote collaborator. The solution is thus lightweight and requires minimal hardware instrumentation on a commodity tablet, hence suitable for nomadic workers' remote collaboration using mobile devices (see Figure 4.4). Second, a user study was conducted showing the effect of unmediated hand gestures compared to a baseline sketch-on-video interface in remote collaboration on physical tasks using mobile devices. The user study showed that unmediated hand gestures really improved user performance and reduced their mental workload in collaboration using mobile devices compared to the sketch-based interface, especially when users were not yet familiar with the task. Specifically, in-air movements of unmediated hand gestures on the device's screen provided collaborators early information about the intended artifacts their counterpart was about to indicate. This thus speeds up their communication instead of waiting to see the sketches performed by touch interaction. Furthermore, unmediated hand gestures aided collaborators to non-verbally communicate complex manipulations that were hard to be expressed by sketches. Third, insights on how users perceived and made use of unmediated hand gestures in the study serve as design implications for future systems supporting workspace awareness on mobile devices. Apart from that, I believe that the MirrorTablet add-on's design combined with suitable visualizations of hand gestures can be also applied for other remote collaborative tasks with involvements of mobile users, such as hub-satellite problem solving using large display-mobile devices setups as in DigiMetaplan. This work is associated with ***Paper C*** in the ***Summary of Articles*** section.

4.2.4 Improving gaze interpretation in remote collaboration

As mentioned earlier, one of the major challenges in HyperCollabSpace is how to convey the satellite worker's attention on certain people and artifacts on the hub side to the hub coworkers. A person's attention can be conveyed via different cues. One can verbally mention what he/she is focusing on, or use pointing gestures to refer to the attended object. However, the most subtle and unobtrusive cue for conveying attention is gaze - a person's looking direction conveyed via the direction of head and eyes [37, 44, 71]. In fact, before verbally describing or pointing at an attended object, people tend to look at it to determine where it is located. Therefore, gaze is considered as a reliable predictor of conversational attention [144, 145] and offers effortless reference to spatial objects [2]. Thus, supporting the hub coworkers to effectively interpret the remote participant's gaze has been explored by previous work to convey users' attention in collaborative systems [27, 38, 124]. Both mediated and unmediated approaches have been studied to support gaze awareness on a shared workspace. Mediated approaches typically use graphical or physical widgets such as circles, arrows or hemispherical displays to represent gaze information, either acquired from eye-tracking data or manually indicated by users [2, 27, 110, 157]. While these approaches can be applied for remote individuals using commodity personal devices like laptop or tablet, such representations might lead to users' distraction due to spatial and representational disparity to the video feed of the remote participants. In contrast, unmediated approaches [**Otsuka-201**, 59, 103], which leverage conveying gaze via presentations of 2D videos or 3D point-clouds of the remote participants, despite being unobtrusive, typically require stationary and specialized hardware configurations, not suitable for traveling scenarios. To address this research gap, the last piece of work included in this thesis proposes GazeLens, a design solution to help hub coworkers interpret the unmediated gaze of a satellite worker conveyed in video conferencing more effectively(see Figure 4.5). I chose to leverage unmediated gaze in video conferencing because this type of media still remains as a dominant form to represent remote participants despite the efforts in developing advanced 3D displays [43, 75, 141]. The reason is that it needs only commodity hardware compared to complex setup required by the latter. Building up on the hub-satellite team settings in DigiMetaplan and HyperCollabSpace, GazeLens also explores a similar scenario commonly found in reality: a satellite worker remotely working with his/her hub coworkers sitting around a table; the satellite worker views the hub's space (people and artifacts on the hub's table) on his personal computer screen and the hub coworkers see the satellite worker in a video feed shown on a computer screen placed at an edge of the hub's table. GazeLens achieves the purpose of improving the hub coworkers' accuracy in interpreting the satellite's gaze by strategically leveraging the entire satellite's screen space to display the hub coworkers and artifacts. More specifically, first GazeLens system captures two video feeds on the hub's space: a panoramic video of hub coworkers by a 360° camera placed at the center of the table and a video of the hub's table from a ceiling-mounted camera. GazeLens combines these two video feeds into an interface on the satellite's screen and strategically uses virtual lenses to guide the satellite worker's attention towards specific areas of their screen, so that hub coworkers can clearly interpret their gaze direction. GazeLens

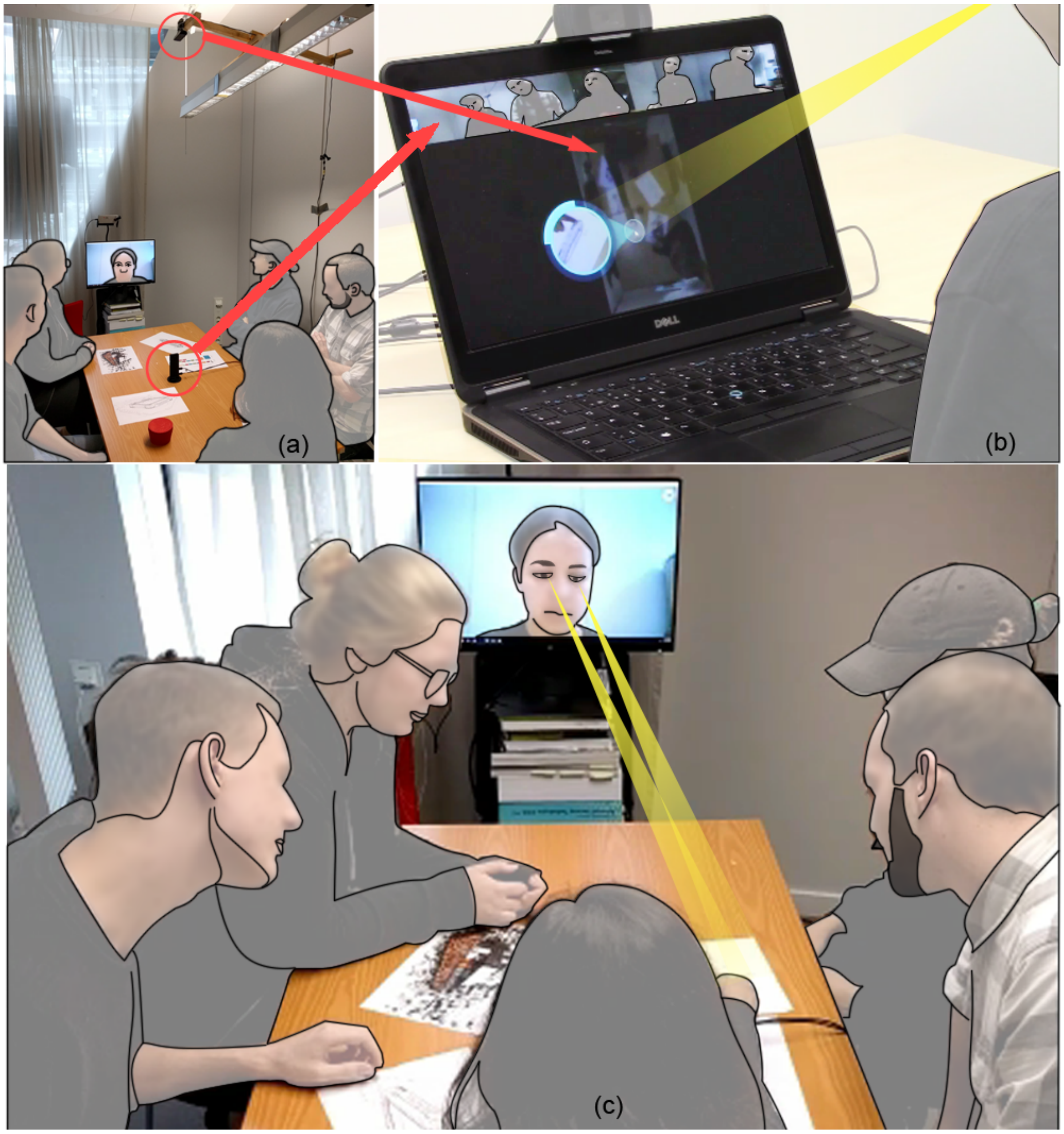

Figure 4.5: GazeLens system. (a) Setup on the hub side with a 360° camera placed on the table to capture hub coworkers and a conventional webcam mounted on the top to capture artifacts on the hub's table. (b) Video feeds from the conventional webcam and 360° camera are separately placed on the remote participant's screen; a virtual lens strategically guiding the remote participant's attention to an attended artifact. (c)Remote participant's gaze, guided by the virtual lens, aligned with the in-focus artifact on the hub's space.

does not require any special hardware instrumentation, but only a conventional laptop or tablet on the satellite worker's side, thus suitable for nomadic settings. On the hub side, it only requires an off-the-shelf conventional webcam and a 360° camera for capturing video feeds. Besides that, this piece of work also provides a user study demonstrating the effectiveness of the proposed systems as well as insights on how screen space of commodity laptops or tablets can be leveraged to better support non-verbal referral communication in remote collaboration. This work is associated with ***Paper D*** that will be described later.

4.3 Summary of articles

This section summarizes contributions and experimental processes of each article included in the thesis. One can easily realize that the order of the papers is aligned with a shifting from mediated to unmediated cues to support workspace awareness in collaboration of distributed teams.

Paper A discusses how workspace awareness supports influence collaboration dynamics of hybrid teams in structured collaborations in terms of role changing and phase shifting. The paper first presents the design and implementation of DigiMetaplan, a digital system that supports facilitated creative problem solving in both collocated and distributed settings. The design of DigiMetaplan is grounded on the practices of Metaplan, a well-established structured brainstorming process in companies and existing systems in the literature. According to users' roles in the team as well as the current phase of the team in the process, DigiMetaplan equips them with different tools to attract their awareness on activities on the brainstorming canvas - the shared workspace. A user study on hybrid teams (3 co-located + 1 remote) shows that such design of DigiMetaplan ensures teams' collaboration consistent with the actual practice of the conventional Metaplan method, suggesting that DigiMetaplan's design effectively supports role changing and phase shifting in a structured collaborative process. Besides that, quality observations also suggest that DigiMetaplan's awareness-support features could enhance a team's collaboration dynamics compared to conventional paper-based setups. This paper sheds light for future systems to consider users' roles and collaboration processes in supporting workspace awareness.

Paper B presents HyperCollabSpace, a concept of a virtual environment that aims to support satellite workers to immersively collaborate with their hub coworkers sitting in a room and working with a shared workspace here, an interactive digital whiteboard. The aim of fostering immersion is achieved through providing a satellite worker with a richer awareness on the presence and activities of their hub coworkers, especially on the shared workspace. A virtual representation of the hub's space was created from a panoramic video of the room, a digital copy of the shared workspace and 3D kinetic avatars of hub co-workers working with the shared workspace. All these elements are positioned so that they are spatially consistent with the actual arrangement of people and artifacts in the physical space, allowing viewers to infer adequately correct targets of hub coworkers' activities, especially on the shared workspace. The satellite worker can view and freely navigate this virtual environment on a commodity tablet, allowing them to observe the hub's space from different perspectives, avoiding cluttered views typically found in conventional video conferencing. Additionally, a write-on-glass mode allows the satellite worker to interact with the shared workspace (e.g. pointing, sketching) while maintaining his/her awareness of all hub coworkers, including those working with the whiteboard and those sitting afar. Potential studies on the concept will be also discussed to inspire future work.

Paper C investigates a design that supports presenting unmediated hand gestures in remote collaboration using mobile devices. To achieve that, the paper first introduces the design of a light-weight and low-cost hardware add-on called MirrorTablet that leverages the form factor of a commodity tablet/smartphone to capture images of a user's hand gesture on and above the device's screen. This design supports the mobility nature of the device. In remote collaboration where coworkers view a shared workspace on their mobile device, overlaying their unmediated hand gestures on the shared workspace can help improve their workspace awareness. A user study comparing MirrorTablet with a sketch-based video conferencing interface was conducted with ten pairs of participants working on a physical task, where an instructor guides a performer to complete a construction. Results show that when participants are not used to the task, unmediated hand gestures of MirrorTablet helped them perform the collaboration faster and caused less perceived workload. Apart from that, participants raised concerns on the cluttered workspace on the mobile device caused by the presence of unmediated hands. How this could be mitigated as well as how the hardware design can be improved in the future will also be discussed.

Paper D examines a design called GazeLens, that employs only off-the-shelf devices to improve hub coworkers interpretation accuracy of a satellite worker's gaze in video communication. Using a 360° camera on the hub side, GazeLens captures a panoramic video of hub coworkers sitting around a table. Artifacts on the table are captured using an RGB camera mounted on the ceiling. On the satellite worker's side, these two video feeds are presented on the screen of a commodity laptop. Virtual lenses were employed to strategically guide the satellite worker's attention to a suitable location on the screen while viewing hub's people and artifacts so that his/her gaze was better aligned with the corresponding targets on the hub side. Two user studies were conducted comparing GazeLens with a conventional video conferencing interface. Results show that GazeLens improves the interpretation accuracy of the hub coworkers on the satellite worker's gaze; and help the hub coworkers better distinguish the satellite worker's gaze on people and artifacts. Early feedback from participants performing a collaborative physical task also suggests positive usability of GazeLens. Finally, further discussions provide insights into how to leverage the screen space of personal computers to improve non-verbal communication in video conferencing.

4.4 Summary of findings

A major finding in DigiMetaplan is that by flexibly offering different levels of access to digitally manipulate a shared workspace according to user roles in a process, a collaborative system can effectively improve a group's global workspace awareness and coordination. Gutwin and Greenberg [49] raised the question on how to balance between the individual's manipulation freedom and the ability to maintain global workspace awareness of a group in collaborative systems. Obviously, when individuals have high levels of manipulation freedom and all of their interactions are visible on the shared workspace, the interface will become distracting. I believe that coupling roles and their corresponding levels of manipulation freedom on the shared workspace can be a helpful tool for less distracting collaborative interfaces while still principally satisfying different users in a collaboration.

Another finding that is apparent in HyperCollabSpace, MirrorTablet and GazeLens is that commodity personal devices like tablet and laptop can still be better leveraged to improve workspace awareness in remote collaboration with the involvement of satellite workers. By creating a virtual environment consisting of both unmediated cues (panoramic video of the hub's space) and mediated representation (3D kinetic avatar), HyperCollabSpace can provide satellite workers an immersive environment where they can perceive interactions of hub coworkers in the workspace spatially consistent with reality. With MirrorTablet, a minimal instrumentation can capture and present unmediated hand gestures on mobile devices, improving workspace awareness in collaboration. A substantial finding from MirrorTablet is that unmediated hand gestures improve workspace awareness in remote collaboration using mobile devices but might also cause occlusion on the workspace due to the small screen size of the devices. Finally, GazeLens shows that the screen space of a commodity laptop can be better leveraged to help hub coworkers better interpret the gaze of a satellite worker, potentially improving non-verbal communication in the collaboration of distributed teams. I believe that these technical findings are worth exploring in the consumer technology landscape where screen-based devices still remain a dominant platform for people to distantly communicate and work with each other, while devices providing richer interaction dimensions such as mixed-reality/virtual-reality headsets still seem to be impractical for work practices of nomadic workers.

Chapter 5

Discussion

In this section, I will first reflect on the overall achievements as well as the remained weaknesses of the thesis. Next, I will attempt to discuss various aspects in supporting workspace awareness in remote collaboration using 2D surfaces. These aspects are what I encountered in the process conducting the pieces of work included in the thesis and seem to be inadequately noticed in the literature. Even though this is only my own reflection based on my observations and literature survey, I hope it will provide the community considerable perspectives to design future collaborative systems.

5.1 Overall reflections on the thesis

Overall, the thesis is a multifaceted endeavour that contributes different technical solutions supporting workspace awareness in remote collaboration using interactive 2D surfaces. The solutions presented here were iteratively developed through user-centered design processes to ensure that they will address the users' actual needs. Several methods ranging from literature review, user observations to interviews were employed to adequately elicit the users' needs and thereby to define design requirements. The implementations of the designs leverage the use of off-the-shelf computers, especially personal and mobile devices, with minimal hardware instrumentation. This makes them promising to be adopted in the near future, especially by nomadic workers who always aim for lightweight and compact hardware setups. The high-fidelity prototypes were also evaluated using various methods to not only validate if the designs could meet the requirements but also to provide insights on their effects on humans' interaction and communication in remote collaboration. As the proposed designs tackle various aspects in human-computer interaction and computer-supported collaborative work, evaluating such systems is complicated, requiring the use of suitable techniques for adequate and appropriate investigations. In this thesis, I employed several empirical methods (e.g. qualitative observations, quantitative evaluations) and data collection techniques (e.g. measuring users' performance, questionnaires, interviews). Those were chosen based on different factors related to the contexts of the designs such as the problems that need to be solved, the existence of baseline systems, the fidelity and robustness of the prototypes and established approaches utilized in the respective studies. In general, the results reported provide insights that were perceived by paper reviewers as interesting contributions to the community; this is an indicator that the chosen methods were appropriate for the designs.

Nevertheless, I believe the evaluations of the systems in the thesis are still far from being complete. Although the thesis succeeded in examining the usability of the proposed designs, it lacks of further investigations on users' experience in using the systems, for example hedonic experience [28]. In fact, understanding if a design is pleasant, comfortable and satisfying to its users is a very important aspect in interaction design research as those are the characteristics that determine whether people like a system or not, once their requirements of usability are sufficiently fulfilled. There also exist some standardized tools for evaluating those metrics that can be employed to improve the studies in this thesis, such as AttrakDiff scores [1], which has been used in the literature [94, 95]. However, performing such investigations on novel interfaces requires careful considerations. Prototypes of novel interfaces especially in academic research are often quite far from their final real-life implementations due to differences in engineering focuses as well as technological availability. Such prototypes are mainly to explore the effect of a novel concept which in many cases cannot be realized using a single off-the-shelf device. Therefore, researchers quite often need to build the prototype by combining different hardware they are currently have in their hands, which makes it way too bulky, less user-friendly and even have much worst performance than what the researchers envision given the availability of suitable hardware and software technologies. It is generally recognized that research lab prototypes cannot reach the same level of engineering sophistication as final products. This gap potentially makes it difficult to reliably evaluate users' hedonic experiences of novel interfaces from research labs. Hedonic experience evaluations using those prototypes would easily lead to feedback that is more negative than what we can receive from polished final products. Thus, even though evaluations on user experiences, especially the hedonic one, are needed, researchers need to consider the aspects to be evaluated that are less influenced by the polishment of the implementation and the maturity of the technologies.

Another limitation of the thesis is the lack of studies in-the-wild on the developed concepts in order to gain more insights on the systems' usability in less controlled environments. In fact, deploying a system in the wild to collect real-world insights is challenging for a couple of reasons. First, such evaluations require a high level of engineering perfection found in final products, which can be difficult with novel lab interfaces built with home-brew instrumentation. Such home-brew instrumentation, such as some 3D-printed hardware add-on or software-based approaches, needs to be much more improved either in term of form-factor designs or performance optimization which sometimes might require specialized human resources to accomplish. Second, polishing a prototype for in-the-wild studies might sometimes require overhead efforts for practical tasks that are not really affordable for an academic research projects. For example, in order to deploy a remote collaboration system, beside polishing the

[1] `http://attrakdiff.de/index-en.html`

proposed interface, it is very important to develop a robust back-end system and reliable protocols for transmitting different types of media (e.g. video, audio, sketches) over different kinds of network connection (instead of only local ones) with a minimal latency. Besides that, it is also crucial to design a configuration interface that allows people to quickly and easily setup a remote collaboration session using the proposed interface without the regular support from the researchers. Having these will ensure that the users will keep using the system in their real-world activities because people always want a computing system to help them live and work more efficiently rather than making them struggle. However, such engineering tasks are typically effort-consuming, but often not regarded as an important part of a research project from scientific-contribution perspectives. Likewise, finding suitable test subjects for such long-run studies is also difficult for many academic research teams. First, such test subjects must be willing to allow the researchers to deploy such the system in their workplace. Second, the test subjects need to perform the activities relevant to the research quite frequently, for example in the case of this thesis, I would have needed to look for teams that often perform remote collaboration on brainstorming, expert-novice support or meetings with involvements of physical artifacts. However, these are particularly challenging due to different workplaces' constrains and people' willingness in the reality.

In short, although in-the-wild user studies are necessary for holistically evaluating a new interface or a new interactive system in HCI, performing such studies is in fact quite challenging for many academic research teams due to limitations of different kinds of resources. In the context of remote collaboration, it is even more difficult as it might require overhead resources that are not affordable by a typical academic research project. Thus, even though on the one side I strongly support running such studies if the team can provide resources for that, on the other side, I believe that research teams could choose other alternatives, such as iteratively running short-term laboratory experiments to investigate, and potentially improve, a design in different aspects with affordable costs.

5.2 Adaptive workspace manipulation ability

One design consideration that can be derived from DigiMetaplan is that reducing the abilities of certain users to digitally manipulate the shared workspace can benefit the group process due to the group's improved workspace awareness. More particularly, as team members cannot directly manipulate the notes, all manipulations must be performed by a single person, the facilitator, helping all team members maintain the same awareness on what is happening on the shared workspace. This finding confirms a suggestion by Gutwin and Greenberg [49] that the more power each individual has in digitally manipulating a shared workspace, the lower the group workspace awareness is. This is because digital manipulation techniques are not bounded by physical constraints. For example, a remote person or a person sitting away from a shared interactive display can easily manipulate content on the display from a distance through their smartphone, which is impossible if they want to do that physically. These digital manipulations can be easily dismissed by the people working on the display (e.g. the faciliator) if they are focusing on another area, thus reducing group workspace awareness. Therefore, the effectiveness in supporting workspace awareness of DigiMetaplan depends upon the design being able to bind the digital interaction techniques, making them more consistent with the physical practices.

As individuals usually tend to want more power to manipulate the shared workspace [49], DigiMetaplan's design obviously leans towards the opposite direction, prioritizing the group's performance rather than the individual preferences. It should be noted that this approach was effective partially thanks to the structure of the original Metaplan method, which facilitates different roles in the process and makes team members more easily accept such restrictions. In a role-relaxed team, where all members have almost the same role, for example in collaborative sketching [14, 128, 155], it would be inappropriate to restrict the manipulation freedom of any of them. Therefore, when considering restricting the digital content manipulation of certain users in a collaborative system, designers should consider the importance of role separation in the collaboration process of the team. Besides that, although DigiMetaplan improved group efficiency overall, certain participants still expressed their wish to have more manipulation freedom on the shared workspace distantly using their mobile device, especially for aiding the facilitator to quickly manipulate a distant object. This would require careful choices from designers in order to balance between group's efficiency and individual needs. Designers can consider offering the facilitator the functionality to instantaneously provide and revoke team members' advanced ability to manipulate the shared workspace. However, such features need to be appropriately unobtrusive, causing insignificant interruptions to the facilitator's workflow while being adequately visible to maintain every team member's workspace awareness. For example, collaborative gestures for distant interaction on large displays like CoReach [93] can be adopted and customized as a tool for a facilitator to quickly and intuitively grant a certain team member temporary access to manipulate a distant object on the shared workspace.

5.3 Temporal aspects of workspace awareness

When it comes to workspace awareness, we typically think about ***at-present*** knowledge about "who", "what" and "where" factors of an event happening on the shared workspace, which means the knowledge of events happening on the shared workspace at the moment a collaborator is concerned. However, through observations on the use of the timeline in DigiMetaplan as well as relevant literature that I have gone through, I would like to raise more design considerations on the historical aspects of workspace awareness - knowledge of "who" ***has done*** "what" at a place ("where") on the shared workspace. Although Gutwin et al. [52] have considered the aspect of "what participants have done" in designing a collaborative learning system, insights into how design features supporting this aspect influence teams' collaboration dynamics have not been explored. Apart from that, the authors also did not take into account the "where" information of the activities in this aspect of workspace awareness. DigiMetaplan, on the other hand, explicitly supports the "what" and "where" aspects as users can see changes in locations and colors of notes or the appearance of hand drawing over time. Due to the role separation in the DigiMetaplan system, the “who” factor was not necessarily accounted for, as all manipulations were performed by the facilitator. However, I believe that the “who” factor is still essential, especially in systems where all users can manipulate the shared workspace. Contradictory to the temporal considerations of DigiMetaplan and Gutwin et al. [52], Prouzeau et al. [120] suggested that in collaborative emergency response systems, historic information on “who” and “where” aspects on the shared workspace are more important to keep track of while the information of “what has been done” can be neglected. I argue that this design choice is suitable for collaborative tasks that focus mainly on the final result (e.g. in emergency-response tasks, the teams only focus only mitigating arising problems) and less on group’s learning. However, in collaborative tasks where the group learning outcomes play an important role (e.g. in brainstorming, the main focus is not the final layout of ideas on the shared display, but the global understandings of all team members on the rationales of decisions made during the process), providing “what” is crucial in the design. Furthermore, temporal workspace awareness cues were also differently presented across systems: While DigiMetaplan’s timeline and similar designs [46, 80] allow users to review a single past snapshot at a time, Prouzeau et al. [120] provide users accumulated up-to-moment traces of collaborators’ activities throughout the progress. I believe that the role of global learning outcomes play a role here. For emergency tasks, collaborators need to know quickly whether an area on the shared workspace has been “touched” by anyone in the team to react appropriately, therefore a view of accumulated traces is beneficial to provide users an overview of collaborators’ visits. For tasks aiming for global understandings of the whole team, viewing a single snapshot at a time is important for all team members to establish equal common grounds on what has happened. Furthermore, in presenting accumulated traces, especially in systems consisting of more than two users, it could be interesting to investigate different representations other than color-coded cues as used by Prouzeau et al. [120].

5.4 Degrees of representational mediation

At the beginning of the thesis, I briefly described two categories of representation forms of workspace awareness cues: mediated and unmediated. However, I believe it would be understandable for a reader to question on whether there are intermediate forms between these two categories; and how to quantify the mediation of the representation of a workspace awareness cue. Reflecting on the designs developed throughout my doctoral study as well as surveyed in the literature, I will discuss these venues based on three main pieces of information - “what”, “who” and “where” as in the framework for workspace awareness in real-time groupware suggested by Gutwin and Greenberge [48].

As mentioned in the “Related work” section, the main difference between mediated and unmediated forms is whether a user needs to spend effort in “decoding” the awareness information conveyed by the cue. However, in order to build future tools for quantifying such users’ “decoding” effort, researchers first need to define what needs to be decoded. As “workspace awareness” was defined as up-to-moment knowledge on “who” is doing “what” and “where” on the shared workspace, I believe that these three types of information can be used to measure users’ effort in different perspectives. “Who” is a measure to quantify how well a design helps its users recognize their collaborators via the cues presented on the shared workspace. For example, a color-coded indicator (e.g. mouse cursor in a shared text editor) may take users more time and higher perceived mental workload to map from its color to the corresponding user it represents compared to using facial images of the user. “What” allows researchers to measure how well a design helps its users understand what their collaborators are doing on the workspace. For example, sketches take more time and cause more confusion than images of remote collaborators’ hand gestures to express and interpret complex manipulations on physical artifacts [78, 87]. Finally, the “where” aspect focuses on quantifying how easily an interface helps users to identify where on the shared workspace their collaborators are working. For example, having a miniature showing the entire workspace of a shared document with rectangles approximately indicating where on the document all collaborators are working on takes less time and probably less physical fatigue to track their coworkers’ current areas of interest compared to interfaces that require them to scroll to look for this information [52]. Considering these pieces of awareness information, researchers can thus break down their study design into quantifying performance of a workspace-awaress support in the aforementioned indicators.

From the design perspective, if designing to support workspace awareness is based on improving users’ performance in “who”, “what” and “where”, there can obviously be different degrees of mediation between unmediated and mediated representation depending on the context. For example, to convey “who”, HyperCollabSpace provides a slightly more abstract representation using 3D avatars with appearance synthesized from physical outlook of users (e.g. height, clothing and hair color) instead of using unmediated images of the remote partners. Similar approaches are only commonly used in commercial VR or MR applications, using pre-defined 3D avatars. Such representations might require more effort from the user to recognize whom the avatar represents than unmediated representations, but less effort than color-coded widgets. On the other hand, this approach also requires more hardware support than color widgets do, but less than unmediated representations like 3D point clouds. Similarly, bodily behaviors of 3D avatars, either captured via depth-sensing camera as in HyperCollabSpace or constructed using kinematics (via handheld remote controllers of VR headset), are better than sketches in conveying what users are doing on the whiteboard, but less accurate than unmediated 3D point clouds due to hardware limitations [151]. Likewise, to convey “where”, there are also different degrees of mediation. For instance, to represent a user pointing at an artifact from a distance, unmediated forms represent as-is actual situations of the user and the artifacts, while an extreme mediated form can even shift user representation next to the artifact [117]. Another intermediate mediation is to keep the as-is distance between the user representation and the artifacts, but adding a pointing ray connecting them [117]. Interestingly, in this case, mediation can improve user performance but reduce the realism of the representation.

In general, by reflecting on different degrees of mediation of workspace awareness cues, their trade-offs as well as how to quantify them, I hope this would serve as an initial point to develop new tools for evaluating the impact of representation forms on team collaboration.

5.5 Occlusions of workspace awareness cues on 2D surfaces

Common approaches for improving workspace awareness are typically overlaying visual information related to collaborators and their activities onto the representation of the shared workspace. However, these approaches can easily lead to occlusions on the workspace due to the spatial conflicts between the artifacts on the workspace and the presented cues. This issue typically happens with cues representing bodily information. Two obvious exemplary cases of this included in this thesis are MirrorTablet and HyperCollabSpace. With MirrorTablet, the unmediated hand gestures were perceived obscuring the images of the physical artifacts. With HyperCollabSpace, the occlusion problem might mainly occur in the write-on-glass mode, where the content on the shared whiteboard could interfere with the 3D kinetic avatar and the panoramic video on the virtual walls behind it. Apparently, this issue can also happen with previous interfaces employing similar metaphors such as CollaBoard[84] and ClearBoard[68]. CollaBoard placed the images of the remote collaborator on the shared display at their corresponding location, which makes their representation obscure the underlying artifacts on the shared workspace. This can be troublesome if the local collaborator needs to take a look at those artifacts. If such cases happen, the remote collaborator might need to move their location to not occlude them. ClearBoard has a similar interference issue as HyperCollabSpace when it could be difficult for a user to perceive expressions of their remote collaborator if there are any artifacts overlaying at the same location on the shared workspace. The cause of such issues is the "collapse" of the physical 3D environment to their 2D representation on a screen, which reduces the depth information of the subjects. In fact, in the physical 3D environment, collaborators using MirrorTablet or CollaBoard do not entirely obscure the artifacts on the screen, thus another co-located collaborator (if any) can easily find another perspective to view the underlying artifacts. Similarly, people working on two sides of a glass in the reality can adjust the focal length of their vision to focus on a subject (i.e. the content on the glass or on their collaborator on the other side), the other thus will be seen in blurred presentations. However, this cannot be achieved in the write-on-glass interface of HyperCollabSpace and ClearBoard where different depths in the physical space are “collapsed” on a single flat surface. This problem was partially solved in the virtual room of HyperCollabSpace where the satellite can freely navigate around the room to find a point of view to the content on the whiteboard that is not occluded by the 3D avatar. However, on the write-on-glass mode, this still remains. A straightforward solution to resolve this occlusion problem is by changing the opacity of the overlaying subject, making them see-through so that users can see the underlying content. However, this may cause confusing interfaces due to the blended content of the different layers. Besides that, such a presentation is typically impossible in the reality (e.g. normal people do not have semi-transparent hands), thus might cause an unpleasant user experience. Nevertheless, thanks to the advancement of eye tracking or head-pose tracking technology, especially using built-in conventional cameras on mobile devices [65, 150], I believe that future systems can mimic the physical world by allowing interfaces to adaptively present contents according to the angle of view of the user. Similar approaches were employed for 2.5D interfaces with simulated depth information [36, 114, 136] or recently used in pseudo-3D image on Facebook [34]. Furthermore, for the write-on-glass interface, eye-tracking data could also be used to simulate humans' foveal and peripheral vision by increasing the sharpness of in-focus subjects and blurring out-of-focus ones, mimicking the changes of human eyes' focal length.

5.6 Making sense of unmediated gaze in remote collaboration

One apparent weakness that can be easily pointed out in GazeLens is the missing evaluation of unmediated gaze supported by our system in realistic remote collaboration settings (e.g. a co-located team having a remote meeting with a satellite worker on physical artifacts, such as physical prototypes of a product). While there have been several works demonstrating the benefits on gaze awareness in remote collaboration [2, 27, 60, 111], gaze cues investigated in these works are in mediated forms, which are less subtle than unmediated gaze. Therefore, it still lacks of knowledge on the effect of unmediated gaze in remote collaboration. In fact, I did attempt to explore this while gathering early user feedback on the usability of the system. However, I realized that people did not rely on gaze in current practices of remote collaboration using video communication. This is interestingly contradictory to my observations on a pilot study, using a simulated remote setting: the instructor was seated at the position of the screen showing his video feed in the actual study with body parts below his chest hidden to the performer. He thus could only provide his instructions either verbally or non-verbally through his gaze. In this simulated distributed setup, I observed that the performer actually relied on the instructor's gaze to aid their communication. This was also confirmed by the performer in the post-study interview. By contrast, in the study with the actual distributed setting, I observed that the performer on the hub side did not pay attention to the video feed of the instructor on the satellite side. The post-study revealed that the performer did not think that he/she could perceive directional cues from the face of the satellite in the video, mainly because he/she did not see the satellite's hand. As a consequence, the instructor had to find ways to verbally communicate the location of an artifact effectively. For example, participants tended to "encode" locations of objects as rows and columns of a matrix, as a satellite referred to an object: "Pick the one at row A, column 1". Another approach is to describe the location of an artifact in relation to the instructor (e.g. "The one closest to you") or fixtures in the hub's space (e.g. "The one closest to the door"). Interestingly, this issue was not explicitly mentioned in the literature. Therefore, in the first submission of the paper of GazeLens, I did report these findings. Obviously, reporting such findings is not conventional for a paper providing a technical contribution as GazeLens because it does not strengthen the contribution of the proposed design. However, my collaborators and I believe that HCI, as a field of science and a design-related domain, should be transparent about the findings and providing multi-facet knowledge for future considerations in the community. Unsurprisingly, we received opposing opinions from reviewers for such findings. On one side, such reported finding was appreciated by researchers who encountered similar issues and suggested that this should be explicitly raised in the community to trigger further studies as well as to design experimental tasks eliminating it. In fact, on the study of ThirdEye [111], one of the works closest to GazeLens regarding the effect of gaze in remote collaboration, the authors had to propose an unnatural communication procedure where the remote participant had to explicitly tell their coworker to look at them before describing and looking at a target. I believe that this is an improvisation to overcome our reported issue, although it does not comply with natural communication in the reality. On the other side, such findings were considered to reduce the contribution of the paper and should be removed. In the end, for the sake of warranting the publication of the paper, such findings were not reported in the latest version, which was published and is included in this thesis. Nevertheless, I still believe that such findings are needed to be voiced to inform the community about such considerations in designing gaze-support collaborative systems as well as in evaluating the effect of gaze on remote collaboration.

Reflecting on the reduced importance of gaze in remote collaboration, I assume that users' reluctance to rely on gaze was caused by poor gaze support of existing video conferencing solution, e.g. Skype typically places a remote person's video in the corner rather than right below the built-in front-facing camera as suggested by Chen [15], making mutual gaze interpretation inaccurate. Besides that, I also suspect that people seeing their remote collaborator being not in the same physical space with them (e.g. the background environment in the video feed is different from their environment) might reduce the sense of presence of the satellite worker on the hub side. It could thus make them believe that their remote coworkers are not sharing the same spatial coordinates with them, thus directional cues of gaze are not reliable. In fact, Banos et al. [7] argued that the form of a media (i.e. the display medium presenting the media) and the media content (e.g. objects, actors, events in the content) are two main factors influencing on the sense of presence. Thus, altering the content of the video feed or the computer screen might increase the perceived sense of presence of the satellite participant, making hub coworkers rely more on his/her gaze. More specifically, it would be interesting to investigate the effect on gaze reliability in response to different types of background environment in the satellite worker's video feed (e.g. actual background environment, a blank background or a simulated background similar to the environment of the collocated team as employed by Cisco CTS-3010 Telepresence system [20]) (see Figure 5.1). Likewise, I'm also concerned about the impact of the device's form factor on the reliability on gaze shown in the remote person's video feed. A common computer screen, typically in a

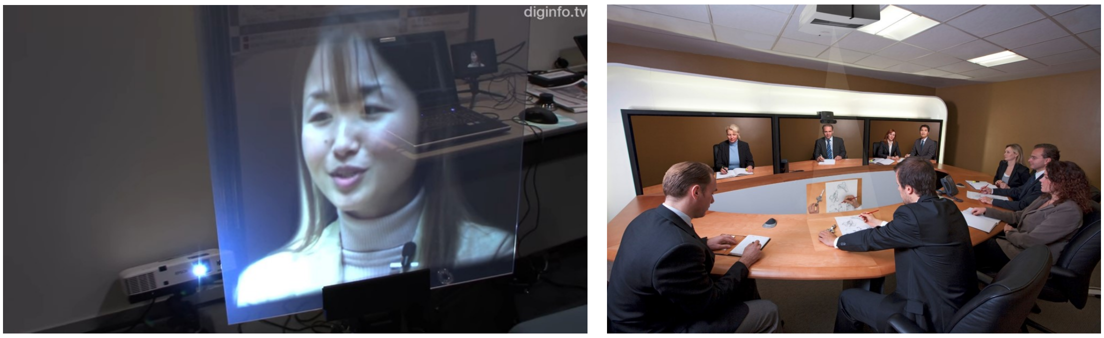

Figure 5.1: Examples of a display medium and a type of media content to be considered in the future to evaluate the effect of unmediated gaze in remote collaboration. Left: using a borderless transparent screen to display the video of the satellite worker with a blank background. Right: the background in the video feed of remote collaborators resembles the physical environment of the local coworkers.

box shape with visible borders, might likely isolate remote people from the physical environment around the screen. Therefore, employing borderless transparent displays [90] (see Figure 5.1) might make the remote person perceived as more blended in the same physical space with local coworkers, potentially improving their trust in spatial coordination conveyed by their gaze. In short, I believe that several properties of the remote collaborator's video feed and display devices can be altered and further studied to improve gaze reliability in remote communication using 2D surfaces.

5.7 Innovating on commodity devices

This reflection on the thesis is more related to my own research mindset that has been shaped throughout my research activities. Personally, I believe that defining a mindset is the most important outcome of any training course as it is the result of the knowledge a person has absorbed through his/her "personal lens".

Human-computer interaction, as a blend of science and design, has been considered as the field where practitioners not only address issues in current interactions between human and computer, but also envision the way people interact with computers in the future, even in a far future. With this mindset, many believe that 2D interactive surfaces, especially mobile devices, are just an intermediate form in the evolution of the computing display medium, thus will soon disappear when VR/MR or holographic technologies mature. Similarly, there is the belief that touch interaction on such surfaces will also vanish and humans will be given back the ability to interact with (digital) artifacts around them entirely using bodily gestures such as grabbing with hands, throwing, pulling, and pushing as they do in the real world. More generally, many researchers often choose to design and study futuristic interfaces and interaction techniques based on assumptions about mature infrastructures supporting them at some point. However, I have been also asking myself a question during my research: *Assuming there will be such future, what will be in between of it and the present? What devices will people use and how will people interact in the intermediate period before* *such a future?* Regarding the aforementioned VR/MR or holographic technologies, there have been apparent problems with them that still remain unsolved such as bulky setups or social unacceptability of their current form factor and interaction modalities, witnessing a slow progress in the development of technical solutions to address them. This dramatically hinders the mass adoption of those technologies as mainstream platforms especially for work purposes. In the context of remote collaborative work, the adoption progress is even slower due to the "technology readiness" aspect, referring to whether all sides of coworkers have sufficient technology infrastructure to carry out the collaboration as expected. If there are infrastructure mismatches among the sides, it might lead to disruptions in their communication, consequentially causing collaboration failures. Therefore, designers and developers are typically reluctant to employ non-commodity devices in developing collaborative systems for reality. Nevertheless, it is always needed to mitigate limitations in user interactions with commodity devices to improve their performance and experience. Thus, I believe that improving user interaction and communication, especially in collaborative systems, should essentially offer transitional technologies, leveraging the use of commodity devices to assure its large-scale acquisition. This mindset has been the driving factor throughout my research, clearly reflected in the contributions. While DigiMetaplan, HyperCollabSpace and GazeLens only require commodity personal computers (laptop, tablet) on the satellite sides and use only off-the-shelf hardware on the hub side (interactive whiteboard, depth-sensing camera, 360° camera), MirrorTablet minimally instrumented a conventional tablet with a lightweight hardware add-on that leverages the tablet's form factor. These are also the two common approaches for innovating on existing technologies that can be found in literature: (1) developing software-based solutions leveraging given devices' form factor and capability; (2) providing minimal instrumentation to extend the interaction space of commodity devices. By enhancing users' capabilities at low cost, these approaches can be quickly and easily taken up by device and gadget manufacturers. Furthermore, I believe that such relatively incremental advancements also act as technology transitions, familiarizing users with novel interaction modalities using their commodity devices before the maturity of specialized consumer hardware. Augmented reality is an example of this. Before AR headsets became commercial gadgets as nowadays, mobile users have long been able to experience augmented reality using their smartphones and tablets, which prepared them a technological background to adopt the headsets. Furthermore, in transitioning from current commodity devices to futuristic platforms, designers can also couple them to offer novel interactions that mitigate the limitations of each single platform. For example, by coupling mobile devices and AR glasses, designers can extend the limited output space of the mobile devices while allowing to use touch interaction for manipulating virtual contents shown on the glasses [47]. This approach not only retains the interaction modality of each device and could also elicit compound approaches. As users' mental models are argued by Craik [24] to be developed through users' interaction with the external world, this approach should thus allow their mental models to inherently shift from the current interaction paradigm to the future one. Furthermore, effects of such transitional compound modalities on users can benefit manufacturers in perfecting their products. Therefore, when a disruptive interaction paradigm is introduced, besides the efforts needed to make it mature, I believe that innovations on existing infrastructure are also comparably important to warrant an inherent shift of paradigms. Furthermore, I hope my reflections developing transitional interactive technologies will offer HCI research another perspective in evaluating contributions in the field as well as in reducing the gap between academic research and HCI practice.

5.8 HCI research with industrial involvement

In this final reflection, I would like to share some insights relevant to my research context. As mentioned in the beginning, my doctoral research started with the MERCO project - a research collaboration between academic and industrial sectors with different outcome expectations based on their business. As one of the primary research contributors, at the end of the project I successfully delivered expected outcomes for both sectors. For industrial partners, the prototype of DigiMetaplan was used as a demonstrative application in the meeting-support ecosystem of AVS Systeme AG (see Figure 5.2). Besides that, Ericsson AB continues further integrating new features into the concept of DigiMetaplan to employ it as a remote collaboration solution at the company. Overall, MERCO was validated as a successful research project by Vinnova [2], the Swedish funding agency. I also fulfilled the expected outcome from academic perspectives with Paper A, which provides HCI scientific knowledge on the effect of workspace awareness on hybrid teams. These are also representatives of typical expected outcomes in research projects with involvements of the two sectors. Hence, to achieve successful results from such projects, it is extremely important for researchers to ensure that their industrial partners will receive expected outcomes (e.g. the prototypes of the research), which were developed based on theoretical insights of the academic researchers. These expected outcomes might be contradictory to each other, due to typical differences in technological assumptions of the two sectors. While academic research tends to look towards the future, assuming on the availability of technologies that still have not existed yet, industrial sectors are typically reluctant to "homebrew" laboratorial technologies due to their high uncertainty to be employed in their businesses. On one side, academics might perceive such differences as limiting their research freedom. On the other hand, I see this as an opportunity to bring academic research closer to the reality.

Colusso et al. [23] surveyed main reasons why academic research is typically not approachable to HCI practitioners. One of the reasons pointed out was that academics tend to provide general theories, which are usually not applicable to realistic design projects that are most of the time context-dependent. However in my experience, our industrial project partners were quite progressive in adopting novel technologies and rather active to attend academic and commercial technology events. Therefore, I believe that when companies are still struggling with a certain problem, it is very likely that academic research is still missing some pieces of solution to address it in their particular context. For that reason, research collaboration

[2] `https://www.vinnova.se/en/`

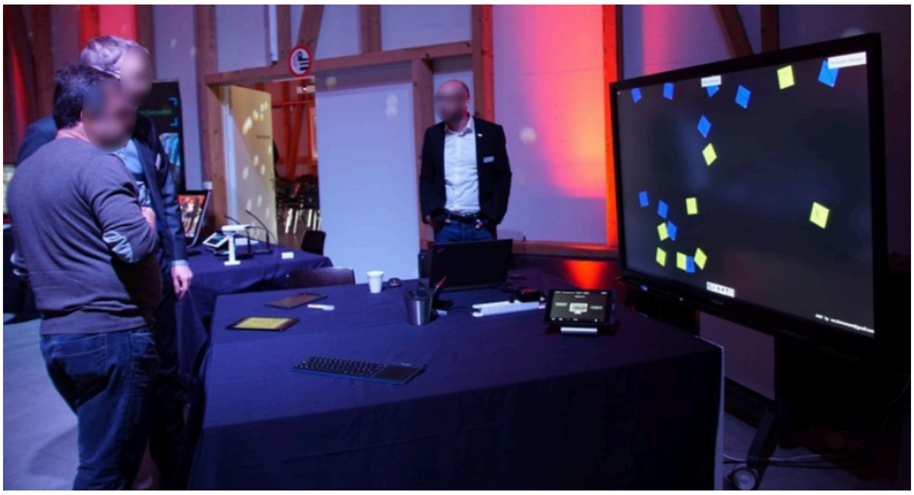

Figure 5.2: DigiMetaplan was showcased by AVS as an application in the meeting-support ecosystem of the company.

between academia and industry would help academics gain more insights into what knowledge is essentially needed by practitioners in their given context, making their design guidelines and recommendations more applicable and actionable. Furthermore, collaborating with the industrial sector also provides academic researchers insights into actual user practices that might be typically dismissed in academic literature, which typically focuses on generalization. For example, in this thesis, I explored creative problem-solving in structured teams, which is quite common in actual practice but not adequately considered in academic research. The studied subjects of HCI are interfaces, which can be used by millions of people all over the world everyday. HCI findings, especially the ones based on commodity hardware, if proven useful, can be easily implemented and released to millions of users in a short time-frame. Thus, making HCI knowledge applicable by practitioners is crucial as it will help the field maintain its pragmatic values that it is supposed to provide. Thus, even though findings in HCI research with involvement of both sectors may be considered niche, they are still needed to advance the field.

On the other direction, academic research also benefits industrial project partners in such collaborations. Research collaborations with academic institutions often have higher levels of technological novelty compared to existing infrastructures in companies. Such novelty also comes with risks and needs to be extensively studied to aid the companies in deciding whether they should integrate them into their business. Managers usually consider well-justified data to make decisions on any new features of a product [23]. However, sometimes, due to the companies' business models, they do not have appropriate research labour to carry out such activities. Thus, collaborations with academia will help them mitigate such issues. Academic researchers often need to comply with scientific standards required by publication venues or academic committees. They thus can provide more reliable evaluations, helping reduce risks in companies' decision-making process.

In short, due to the pragmatic nature of HCI, the existing gap between academic research and actual design practices needs to be reduced. To achieve that, collaborative research between academia and industry should be fostered to improve knowledge exchange between the two sectors and to harmonize theoretical and pragmatic aspects in the HCI community.

Chapter 6

Outlook: Virtually Extendable Slate

In this part, I would like to provide my vision on how the designs included in this thesis can be synthesized in a system that can provide rich workspace awareness to support nomadic workers to efficiently collaborate with their peers in the future.

6.1 Motivation

This thesis has introduced interface design approaches that leverage non-verbal cues such as hand gestures or gaze to improve workspace awareness in remote collaboration using personal computers such as laptops and tablets. However, despite having powerful features, such personal computers encounter certain limitations that cannot be sufficiently addressed by software-based solutions only. For example, even though GazeLens could improve gaze interpretation in video conferencing by leveraging the entire space of a personal computer's screen, the relatively small screen size makes it difficult to align the satellite worker's gaze with targets in a much larger physical space on the hub's side. For instance, with hub coworkers sitting close to the two sides of the hub's screen, the satellite worker's gaze conveyed through GazeLens could not be interpreted correctly by hub coworkers. This is because the field of view of the satellite worker on his/her screen is much narrower than the field of view formed by the hub's screen and the hub's coworkers sitting closest to its sides. This mismatch between the shared workspace and the satellite worker's screen size might also cause similar problems in conveying the satellite worker's gaze in other remote collaboration settings, for example in the setting where a satellite worker uses a laptop or a tablet to collaborate with remote coworkers who are working on a large display. In remote collaboration on large displays (e.g. data analysis, emergency response, crisis management, etc.), settings typically assume symmetric hardware setups, i.e. there are large displays on all sides of the collaboration. Although such settings could convey the remote collaborator's gaze effectively due to the consistent presentations of the shared workspace on all sides [4, 5], they could hinder the participation of satellite workers, who typically do not want to be constrained by such stationary hardware setups.

To improve the interpretation of the satellite worker's gaze in aforementioned asymmetric settings, it is necessary to have a mobile solution to extend the satellite worker's display. To achieve this, I envision a concept called *VXSlate* (***V***irtually e***X***tendable ***Slate***) that utilizes an MR/VR headset to virtually extend the screen space of a personal computer. Even though MR/VR headsets are currently not commodity devices, the continuous improvement efforts of manufacturers promise off-the-shelf gadgets of these technologies in the near future. Besides that, as these devices can allow any individual to work with high-resolution and large-scale contents without the need of bulky physical large displays, they could be massively adopted by nomadic workers when the technologies mature, making asymmetric mobile-large display remote collaboration settings more efficient. However, these technologies also suffer certain limitations related to their interaction modalities. One of the most obvious limitations of users' interactions when wearing MR/VR headsets is that users currently need to rely either on handheld remote controllers or in-air gestures, which can lead to fatigue and Gorilla arm effects [63, 132] in long-term use due to the lack of sufficient counter-forces as when interacting with physical objects. In the context of remote collaboration, the current form factor of those headsets makes it difficult to capture unmediated facial expressions of the satellite worker to be displayed on the remote side. Thus, in remote collaboration with collaborators wearing these headsets, remote collaborators are still typically presented by 3D avatars, which can reduce the personal engagement between collaborators due to their low realism compared to the collaborators' actual images. By coupling a commodity tablet with an MR/VR headset, this might be mitigated. A tablet can offer an interactive surface where the user can perform touch interaction to manipulate virtual objects seen through the headset, avoiding the Gorilla arm effects. Besides that, while holding and interacting with the tablet, its built-in front-facing camera can capture the facial expressions of the user to be shown on the remote side. By strategically placing the virtual content seen through the headset according to the orientation of the tablet, this combination can also effectively convey the user's gaze on the shared workspace. In the next section, I will provide the detailed design of the VXSlate concept and how it can be integrated into remote collaboration.

6.2 VXSlate design

Using VXSlate, a satellite worker will view a virtual replica of his/her collaborators' large display (the shared workspace) in a virtual environment seen through a VR/MR headset that he/she is wearing. This should provide the user high-resolution images of the shared workspace using a compact hardware setup. The satellite worker interacts with the shared workspace via touch interaction on a mobile device (here, a tablet) held in his/her hands. This approach should reduce the Gorilla arm effect, thanks to the counter force the satellite worker's hand will receive from the tablet's screen during the interaction. However, VXSlate does not map the screen space of the tablet to the entire shared workspace as it might lead to imprecise interaction when mapping a small space to a much larger one. Instead, VXSlate maps the tablet's screen space onto a relatively small portion of the shared workspace, expectedly

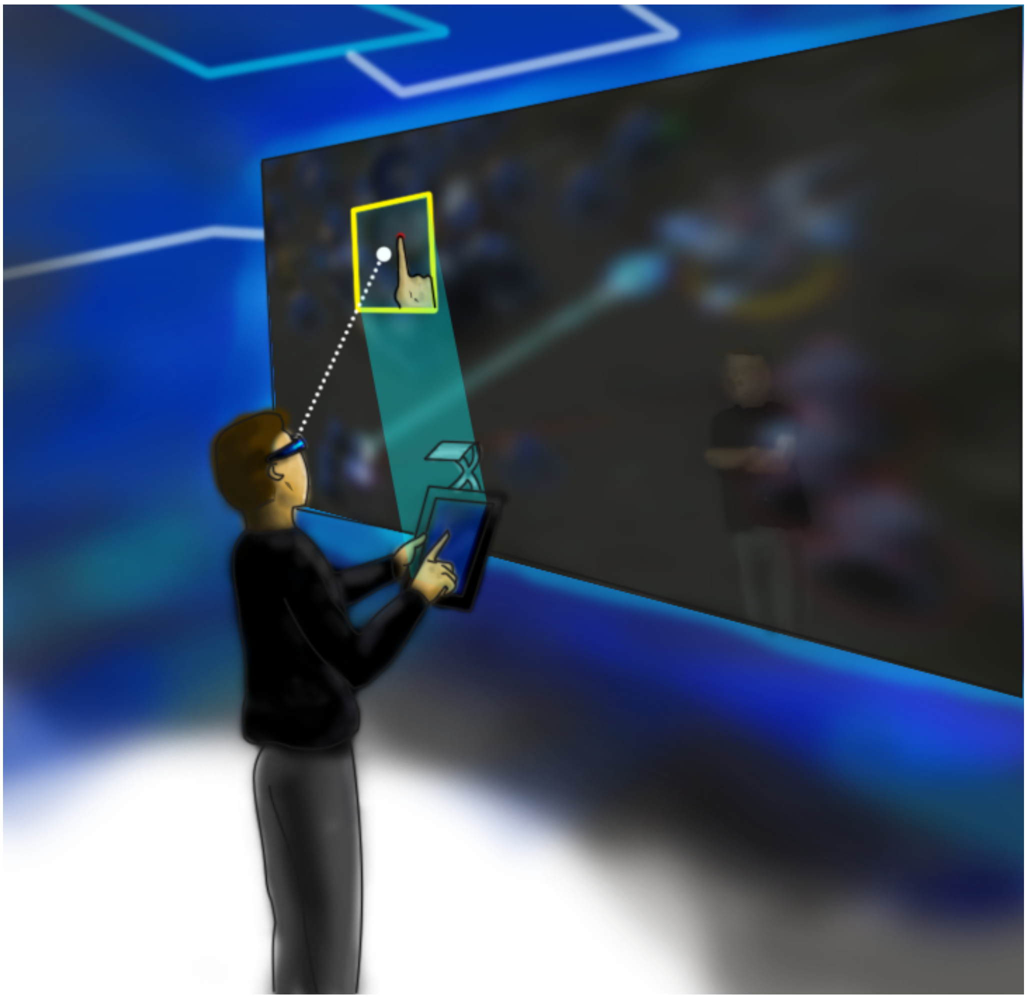

Figure 6.1: VXSlate concept: The satellite worker views a large display virtually through a VR/MR headset, interacts with it using touch interactions on a tablet. Touch interactions on the tablet will be mapped onto a virtual replica, called virtual pad (yellow rectangle), which can be positioned by the satellite worker's gaze (the dashed white line ended up at the white circle in the center of the virtual pad). Touch points on the tablet will be accordingly highlighted by red circles on the virtual pad. Images of the satellite workers' are captured by the MirrorTablet hardware mounted on the tablet and overlaid on the virtual pad.

providing more precise interaction due to less asymmetric mapping of interaction spaces. This interactive small portion of the shared workspace will be briefly referred to as "virtual pad" because it can be considered as the virtual presentation of the tablet that the user is holding. The virtual pad is depicted as a rectangular area on the shared workspace having the same dimension ratio (but might have different scales) as the tablet screen (illustrated as the yellow rectangle in Figure 6.1), where the satellite worker's touch interaction on the tablet screen will be mapped to. With this, the satellite worker can perform various manipulations on the shared workspace within the virtual pad using single- or multi-touch gestures such as object translation, scaling, rotating and even sketching.

Using the aforementioned approach, in order to allow the satellite worker to interact with the entire shared workspace, he/she needs to be able to change the location of the virtual pad on the entire shared workspace effectively. Navigation of the virtual pad should be performed seamlessly, causing as little interruption as possible to his/her workflow and can be achieved at a relatively fine-grained level. To achieve this, VXSlate mainly facilitates the satellite worker' gaze combined with his/her touch interaction on the tablet. Gaze has been shown to be a primary indicator of a person's intention and coordination [106, 66]. When a person is about to interact with a subject, he/she will gaze at it before making the action [115]. Leveraging this, VXSlate couples the position of the virtual pad on the shared workspace with the satellite worker's gaze. When the satellite worker looks at a location on the shared workspace, the virtual pad will be positioned there, with its center at the intersection between the user's gaze and the shared workspace (illustrated as the white circle at the center of the virtual pad in Figure 6.1). This can be easily achieved thanks to the gaze tracking feature available on many consumer VR/MR headsets. At the moment, most headsets provide coarse gaze information of the wearer, conveyed through head-pose tracking. However, current technology developments [1] [2] promise eye-level gaze tracking to be soon available on consumer headsets, supporting fine-grained gaze tracking ability. As human gaze is typically sacchading, VXSlate adopts an approach for gaze interaction on large displays proposed by Turner et al. [142] to stabilize the virtual pad. More specifically, the virtual pad's position will only change when the satellite worker's gaze is outside of it, otherwise it remains unchanged. Although using gaze can be an efficient approach for the satellite worker to quickly shift his/her attention between areas that are far away from each other, using gaze for precise positioning can be tedious. The reason is that gaze fixation is quite sensitive to movement, thus forcing the user to precisely shift his/her gaze within a small distance can be physically demanding. VXSlate complements this by allowing the satellite worker to move the virtual pad around its current position by using touch gestures on the tablet, while his/her gaze is still within the virtual pad. More specifically, the user can perform a two-finger translation gesture to manually shift the position of the virtual pad. This gesture can be differentiated from one-finger translation which is commonly used to move contents within a workspace, while still being sufficiently similar to the latter to be remembered easily. Furthermore, VXSlate also allows the satellite worker to enlarge or downsize the virtual pad in order to expand or shrink the interaction area for quicker or more precise manipulations, using touch gestures on the tablet. The satellite worker can hold the thumb finger of the hand gripping the tablet as an anchor and move two fingers of the other hand to linearly resize the virtual pad. This gesture can be distinguished from two-finger pinching gestures, typically used for scaling contents on the screen; and still possesses certain similarities to be learnt easily.

VXSlate allows a satellite worker to distantly work on a large shared workspace through his/her tablet. In this way, the satellite worker's attention will focus on

[1] `https://www.cnet.com/news/the-eye-tracking-htc-vive-pro-eye-is-a-sign-of-vr-to-come/`

[2] `https://imotions.com/biosensor/eye-tracking-vr/`

the shared workspace while the tablet will act as a touchpad. Due to his/her hand posture while holding the tablet, which is often rather low to avoid fatigue, hand gestures on the tablet are not always visible within the user's field of view while looking at the shared workspace. This implies that VXSlate needs to provide the satellite worker an adequate awareness of his/her hand position on the tablet to avoid incorrect or unintentional touch interactions. Although the tablet's screen can capture the satellite worker's touches, it cannot capture his/her hands when hovering on it. For object manipulations such as translation, scaling or rotating. this lack of visualization of a hovering hand might not be significantly problematic as the satellite worker can fix mistakes rather easily. However, for other manipulations such as precise selections or sketching where mistakes could require a significant effort to mend, that issue might cause large mental and physical workload. Therefore, to support the satellite worker to more easily locate their hand on the interaction space, I believe that it is important to capture his/her hovering hand gesture above the tablet and present it on the virtual pad accordingly. To achieve this while maintaining a compact, lightweight and mobile hardware setup, the MirrorTablet hardware can be attached to the tablet. As MirrorTablet can only capture 2D images of the hand, but not the depth information, VXSlate shows highlighters, in shapes of colored circles, at touch locations to allow the satellite worker to better differentiate between hovering and touching hands. Finally, hub coworkers of the satellite worker can be represented using the content-over-video approach of ClearBoard so that the satellite worker can be aware of different spatial communication cues of them, such as gaze or pointing gestures.

In short, to support a satellite worker to work efficiently on a high-resolution large display using a mobile hardware setup, VXSlate combines gaze tracking of a VR/MR headset with touch interactions on a commodity mobile device to distantly manipulate virtual contents using small-scaled multi-touch gestures. This interaction technique is expected to result in higher manipulation accuracy and to cause less fatigue to the satellite worker's arms. These advantages promise the satellite worker the ability to perform various tasks that need high-resolution large displays to be accomplished efficiently due to complex manipulations such as object translation, scaling, rotation or sketching. Such tasks can be found in different scenarios such as crisis management [19, 121], data exploration[123] and imaging diagnosis [122]. Moreover, the design of MirrorTablet can acquire not only images of the satellite worker's hands but also his/her face. Thus, VXSlate can also capture various non-verbal communication cues such as hand gestures, gaze or facial expressions. These are important expressive factors for tele-communication, making VXSlate beneficial for remote collaboration. The next section describes the concept of a collaborative system that makes use of VXSlate.

6.3 Integrating VXSlate into remote collaboration

When a satellite worker in VXSlate holds a tablet, his/her face can be captured by the tablet's built-in camera and shown on the side of their remote peers, for example the hub's coworkers. First, this provides the hub's coworkers unmediated facial

Figure 6.2: Integrating a satellite worker using VXSlate into remote collaboration on a large display with a hub coworker. The video feed consisting of the face of the satellite worker can be captured by the MirrorTablet hardware and positioned on the hub's large display so that his gaze will be aligned with the targeted artifact. Images of his hand can also be captured by the MirrorTablet hardware and overlaid on the large display at the area of interaction. An animated robot arm visualization is used to link the satellite worker's face video feed and his hand. The content-over-video metaphor of Clearboard [68] is used to mitigate occlusions of the video feeds to artifacts on the shared workspace.

expressions of the satellite worker, warranting representational realism. Second, the direction of his/her gaze while looking at the shared workspace, in relation to the camera, can be used to convey his/her visual attention, using a similar mechanism to the one of GazeLens. By knowing the looking gaze direction of the satellite worker on the shared workspace as well as his/her position in the virtual environment, the video feed of him/her, captured by the tablet's camera, can be strategically positioned on the physical large display on the hub side so that his/her gaze in the video feed will be aligned with the targeted artifact (as seen in Figure 6.2). In this way, the satellite worker's gaze conveyed through VXSlate should be more accurate than through the GazeLens interface, as the satellite worker's field of view in the former is much larger than in the latter. Besides that, the relative position of the tablet to the satellite worker's head can also be tracked (for example, using light-weight technologies like TrackCap [100]), to better calculate the position of the video feed, providing more correct gaze interpretation. Apart from that, images of the satellite worker's hand captured by the MirrorTablet hardware can also be overlaid on the shared workspace on the hub's side at the corresponding area. This can help hub coworkers be better aware and comprehend what the satellite worker is doing on the shared workspace.

Although the MirrorTablet hardware allows VXSlate to present unmediated facial cues and hand gestures, those cues are separated due to the limitations of the tablet's camera. Figure 6.2 illustrates this issue where only the upper part of the body and parts of the hands on the tablet screen can be captured and displayed separately.

This fragmented representation of the satellite worker might cause additional efforts of the hub coworkers to interpret his/her spatial attentions as they need to find a link between two dispersed visual components. Therefore, to reduce the arising mental effort, visual links between the two components are needed to help the hub coworkers to explicitly understand the connection [146]. While some previous work chose to distort a body part of the satellite worker [61], this might lead to unpleasant experiences due to the uncanny valley problem [134], where the appearance of a representation which looks almost, but not exactly, like real human beings, causes strangely familiar, eerie feelings and revulsion to the viewer [67]. Thus, the visual link between the two visual video feeds of the satellite worker can be in a shape that does not resemble a human appearance, such as an animated robot arm as illustrated in Figure 6.2. I hypothesize that this cyborg-like representation not only helps hub coworkers effectively maintain their awareness on the satellite worker's activities on the shared workspace, but also provides more joyful experiences of the collaboration.

Moreover, I believe that VXSlate is a complement to expand the interaction space, and further how a nomadic worker utilizes a mobile tablet for remote collaboration. Depending on the task, the nomadic worker can choose to use the setup that is the most convenient to his/her context. For example, when the task does not need a large shared workspace, for example in document co-editing or pair programming, the shared workspace can simply be shown on the tablet's screen. Otherwise, when there are several artifacts on the shared workspace such as in structuring pages of a document or discussing different design concepts, the nomadic worker can choose to enlarge the shared workspace and he/she can fully view it through his/her VR/MR headset. Likewise, the practices of using multiple displays in an office context can also be digitized by VXSlate where the nomadic worker can work with the main window on the tablet screen and arrange other windows virtually around. I believe that this will significantly improve the work efficiency compared to being limited to only a tablet which is difficult to work with multiple windows due to the limited screen size. Similar mechanisms in leveraging gaze as in GazeLens as well as in capturing hand images using MirrorTablet hardware can be utilized to let remote collaborators know which window the nomadic worker is currently working on or looking at. All of these put together open a future where nomadic workers would almost have no limit to perform remote collaboration anywhere in any tasks with a wide range of complexities. Most notably, that can be done using a remarkably compact hardware setup that all can even easily fit into a small backpack.

Chapter 7

Conclusion

The pieces of work included in this thesis provide different contributions, targeting various workspace awareness aspects in remote collaboration. However, to provide readers concise take-away messages, the contributions can be regarded from the following high-level perspectives.

7.1 Implications for HCI practitioners

This thesis provides multiple insights into the effects of workspace awareness cues on remote collaboration, that can be considered by HCI practitioners (designers, researchers) in designing future systems or for further studies.

In distributed teams consisting of multiple roles and following structured collaboration processes, it is important to provide team members awareness tools according to their roles and depending on current activities in the process. As shown in one of the studies in this thesis, using tools designed following this consideration, participants behaved consistently with their expected behaviors in conventional settings, regardless of their limited experience with the used methods. Contrasted to unexpected issues in the literature caused by inappropriate individual accessibility to workspace awareness tools, this can be considered as a solution for collaboration dynamics of distributed teams. Nevertheless, as this design consideration is inspired by structured collaborative methods that prioritize a team's global performance over individual flexibility, designers should take into account the desired outcomes of the team's collaboration for appropriate application.

In certain contexts, better workspace awareness might not outweigh its cost to user experience. As demonstrated, although unmediated hand gestures were beneficial in remote collaboration on familiar tasks, their occlusive representation on mobile devices could lead to users' dislike. This suggests two design considerations: First, on small-size displays, it is important to foster the use of mediation that offers behavior-consistent and less obscuring representation than unmediated forms. Second, lightweight and simple workspace awareness supports might be sufficient for users with high levels of familiarity to the task and when the sense of presence is not critical to the collaborators.

Instead of requiring additional display hardware as commonly found in the literature, the thesis demonstrates that the screen space of a conventional laptop can be leveraged to improve unmediated gaze interpretation in remote collaboration. This approach can be also leveraged in other shared workspace setups rather than horizontal tables as in our study. For example, combining the spatial metaphor of GazeLens with suitable positioning of video feeds, designers can convey the unmediated gaze of a mobile user towards artifacts on a large display in an asymmetric collaboration setting between an individual using a personal computer and a team working with a large display. Considering this design insight, designers can provide lightweight approaches for facilitating gaze awareness in remote collaboration.

Panoramic videos are a potential medium that promises different applications to support workspace awareness. First, the thesis demonstrated how a panoramic video can be used as a low-cost approach to construct a virtual environment that is spatially consistent with its corresponding physical space, thus allowing remote people to interpret activities of their co-workers in relation to the shared workspace. Besides that, the thesis also showed that a suitable placement and adaptive resegmentation of a panoramic video, leveraging the form factor of the display device, can improve gaze interpretation in video conferencing.

The thesis proposed concepts and prototypes of systems for supporting workspace awareness. Although a number of studies have been conducted to investigate certain venues, I believe that the described systems could serve as platforms for further studies on other aspects such as users' attentions in different types of collaboration, effects of unmediated gaze in collaborative tasks or effects of hand visualizations on small-screen devices.

7.2 Implications for users, manufacturers and developers

All designs presented in this thesis show that commodity devices still possess significant potential to be further exploited to support remote collaboration. HCI, as a field of science, has been argued to focus on the future, designing interfaces with assumptions on the availability of currently non-existing technologies. However, that may also hinder the transfer of academic research to HCI practices, due to the mismatch between hardware infrastructures of the two sectors. Therefore, the designs leveraging off-the-shelf hardware as presented in this thesis could assure more approachable research results to gadget manufacturers. It should be easy for manufacturers to integrate those results into their products, primarily to set out their competitive strengths on the market as their products may offer better interactions and user experience. Furthermore, such adoptions will help users get acquainted to forthcoming interaction modalities, supporting a seamless paradigm shift towards the future, when specialized hardware matures.

To better foster workspace awareness in remote collaboration, awareness supports need to be better integrated into computing ecosystems. The problem with gaze unawareness in video conferencing shows that established design considerations such as leveraging video feed positioning to correct conveyed gaze are still poorly employed in commercial products. From the perspective of developers, this can be understood because such design considerations can lead to interference with design principles of interfaces on commodity computers. To address this, more efforts both from research and production parties are needed to develop new interface standards that allow better adoptions of research outcomes, effectively changing users' habits and thus, improving communication in remote collaboration.

Bibliography

[1] Marilyn Jager Adams, Yvette J Tenney, and Richard W Pew. "Situation awareness and the cognitive management of complex systems". In: *Human factors* 37.1 (1995), pp. 85–104.

[2] Deepak Akkil and Poika Isokoski. "I see what you see: gaze awareness in mobile video collaboration". In: *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*. ACM, 2018. ISBN: 978-1-4503-5706-7. DOI: 10.1145/3204493.3204542.

[3] Deepak Akkil, Jobin Mathew James, Poika Isokoski, and Jari Kangas. "GazeTorch: Enabling gaze awareness in collaborative physical tasks". In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM. 2016, pp. 1151–1158.

[4] Ignacio Avellino, Cédric Fleury, and Michel Beaudouin-Lafon. "Accuracy of deictic gestures to support telepresence on wall-sized displays". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM. 2015, pp. 2393–2396.

[5] Ignacio Avellino, Cédric Fleury, Wendy E Mackay, and Michel Beaudouin-Lafon. "Camray: Camera arrays support remote collaboration on wall-sized displays". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM. 2017, pp. 6718–6729.

[6] Mahdi Azmandian, Mark Hancock, Hrvoje Benko, Eyal Ofek, and Andrew D Wilson. "Haptic retargeting: Dynamic repurposing of passive haptics for enhanced virtual reality experiences". In: *Proceedings of the 2016 chi conference on human factors in computing systems*. ACM. 2016, pp. 1968–1979.

[7] Rosa Marıa Baños, Cristina Botella, Mariano Alcañiz, Vıctor Liaño, Belén Guerrero, and Beatriz Rey. "Immersion and emotion: their impact on the sense of presence". In: *Cyberpsychology & behavior* 7.6 (2004), pp. 734–741.

[8] Hrvoje Benko and Steven Feiner. "Balloon selection: A multi-finger technique for accurate low-fatigue 3d selection". In: *2007 IEEE Symposium on 3D User Interfaces*. IEEE. 2007.

[9] Jacob T Biehl, Daniel Avrahami, and Anthony Dunnigan. "Not really there: Understanding embodied communication affordances in team perception and participation". In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM. 2015, pp. 1567–1575.

[10] Sebastian Boring, Marko Jurmu, and Andreas Butz. "Scroll, tilt or move it: using mobile phones to continuously control pointers on large public displays". In: *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest*. Melbourne, Australia: ACM, 2009, pp. 161–168. ISBN: 978-1-60558-854-4. DOI: 10.1145/1738826.1738853.

[11] Alan Bourning and Michael Travers. "Two approaches to casual interaction over computer and video networks". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1991, pp. 13–19. ISBN: 0-89791-383-3. DOI: 10.1145/108844.108847.

[12] John Brooke et al. "SUS-A quick and dirty usability scale". In: *Usability evaluation in industry* 189.194 (1996), pp. 4–7.

[13] Frederik Brudy, Joshua Kevin Budiman, Steven Houben, and Nicolai Marquardt. "Investigating the role of an overview device in multi-device collaboration". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. 2018, p. 300.

[14] Senthil Chandrasegaran, Sriram Karthik Badam, Zhenpeng Zhao, Niklas Elmqvist, Lorraine Kisselburgh, and Karthik Ramani. "Collaborative sketching with skWiki: a case study". In: *ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers. 2014, V007T07A041–V007T07A041.

[15] Milton Chen. "Leveraging the asymmetric sensitivity of eye contact for videoconference". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2002, pp. 49–56.

[16] Sicheng Chen, Miao Chen, Andreas Kunz, Asim Evren Yantaç, Mathias Bergmark, Anders Sundin, and Morten Fjeld. "SEMarbeta: mobile sketch-gesture-video remote support for car drivers". In: *Proceedings of the 4th Augmented Human International Conference*. ACM, 2013, pp. 69–76. ISBN: 978-1-4503-1904-1. DOI: 10.1145/2459236.2459249.

[17] Lung-Pan Cheng, Eyal Ofek, Christian Holz, Hrvoje Benko, and Andrew D. Wilson. "Sparse haptic proxy: Touch feedback in virtual environments using a general passive prop". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2017, pp. 3718–3728. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025753.

[18] Inrak Choi, Eyal Ofek, Hrvoje Benko, Mike Sinclair, and Christian Holz. "Claw: A multifunctional handheld haptic controller for grasping, touching, and triggering in virtual reality". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. 2018, p. 654.

[19] Apoorve Chokshi, Teddy Seyed, Francisco Marinho Rodrigues, and Frank Maurer. "ePlan multi-surface: A multi-surface environment for emergency response planning exercises". In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*. ACM. 2014, pp. 219–228.

[20] Cisco. *Cisco TelePresence System 3010 Data Sheet*. URL: `https://www.cisco.com/c/en/us/products/collateral/collaboration-endpoints/telepresence-system-3010/data_sheet_c78-579689.html`. 2013.

[21] Herbert H. Clark. "Using language". In: *Cambridge University Press: Cambridge*. 1996, pp. 274–296.

[22] François Coldefy and Stphane Louis-dit-Picard. "DigiTable: an interactive multiuser table for collocated and remote collaboration enabling remote gesture visualization". In: *2007 IEEE conference on computer vision and pattern recognition*. IEEE. 2007, pp. 1–8.

[23] Lucas Colusso, Cynthia L Bennett, Gary Hsieh, and Sean A Munson. "Translational resources: Reducing the gap between academic research and HCI practice". In: *Proceedings of the 2017 Conference on Designing Interactive Systems*. ACM. 2017, pp. 957–968.

[24] Kenneth James Williams Craik. *The nature of explanation*. Vol. 445. CUP Archive, 1967.

[25] Ross Cutler, Yong Rui, Anoop Gupta, Jonathan J Cadiz, Ivan Tashev, Liwei He, Alex Colburn, Zhengyou Zhang, Zicheng Liu, and Steve Silverberg. "Distributed meetings: A meeting capture and broadcasting system". In: *Proceedings of the tenth ACM international conference on Multimedia*. ACM. 2002, pp. 503–512.

[26] Sarah D'Angelo and Andrew Begel. "Improving communication between pair programmers using shared gaze awareness". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2016, pp. 6245–6290. ISBN: 978-1-4503-4655-9. DOI: `10.1145/3025453.3025573`.

[27] Sarah D'Angelo and Darren Gergle. "Gazed and confused: Understanding and designing shared gaze for remote collaboration". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2016, pp. 2492–2496. ISBN: 978-1-4503-3362-7. DOI: `10.1145/2858036.2858499`.

[28] Sarah Diefenbach, Nina Kolb, and Marc Hassenzahl. "The'hedonic'in human-computer interaction: history, contributions, and future research directions". In: *Proceedings of the 2014 conference on Designing interactive systems*. ACM. 2014, pp. 305–314.

[29] 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*. ACM, 2014, pp. 229–238. ISBN: 978-1-4503-2587-5. DOI: `10.1145/2669485.2669517`.

[30] Paul Dourish and Victoria Bellotti. "Awareness and coordination in shared workspaces". In: *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*. ACM, 1992, pp. 107–114. ISBN: 0-89791-542-9. DOI: `10.1145/143457.143468`.

[31] Nathan J. Emery. "The eyes have it: the neuroethology, function and evolution of social gaze". In: *Neuroscience & Biobehavioral Reviews* 24.6 (2000), pp. 581–604.

[32] M.R. Ensley. "Toward a theory of situation awareness". In: *Human factors*. 1995, pp. 32–64.

[33] Katherine M Everitt, Scott R Klemmer, Robert Lee, and James A Landay. "Two worlds apart: bridging the gap between physical and virtual media for distributed design collaboration". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2003, pp. 553–560.

[34] Facebook360. *3D Photos Now Rolling out on Facebook and in VR*. URL: `https://facebook360.fb.com/2018/10/11/3d-photos-now-rolling-out-on-facebook-and-in-vr/`. 2018.

[35] Interaction Design Foundation. *What is User Centered Design?* URL: `https://www.interaction-design.org/literature/topics/user-centered-design`. 2013.

[36] Jérémie Francone and Laurence Nigay. "Using the user's point of view for interaction on mobile devices". In: *Proceedings of the 23rd Conference on l'Interaction Homme-Machine*. ACM. 2011, p. 4.

[37] Alexandra Frischen, Andrew P Bayliss, and Steven P Tipper. "Gaze cueing of attention: visual attention, social cognition, and individual differences." In: *Psychological bulletin* 133.4 (2007), p. 694.

[38] Susan R Fussell, Leslie D Setlock, and Elizabeth M Parker. "Where do helpers look?: gaze targets during collaborative physical tasks". In: *CHI'03 Extended Abstracts on Human Factors in Computing Systems*. ACM. 2003, pp. 768–769.

[39] Susan R. Fussell, Leslie D. Setlock, Jie Yang, Jiazhi Ou, Elizabeth Mauer, and Adam DI Kramer. "Gestures over video streams to support remote collaboration on physical tasks". In: *Human-Computer Interaction* 19.3 (Nov. 2004), pp. 273–309.

[40] Steffen Gauglitz, Benjamin Nuernberger, Matthew Turk, and Tobias Höllerer. "World-stabilized annotations and virtual scene navigation for remote collaboration". In: *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM, 2014, pp. 449–459. ISBN: 978-1-4503-3069-5. DOI: `10.1145/2642918.2647372`.

[41] Anthony Giddens. *The constitution of society: Outline of the theory of structuration*. Vol. 349. Univ of California Press, 1986.

[42] R.D. Gilson. "Introduction to the special issue on situation awareness". In: *Human factors*. 1995, pp. 3–4.

[43] Daniel Gotsch, Xujing Zhang, Timothy Merritt, and Roel Vertegaal. "TeleHuman2: A Cylindrical Light Field Teleconferencing System for Life-size 3D Human Telepresence." In: *CHI*. 2018, p. 522.

[44] Reiko Graham and Kevin S LaBar. "Neurocognitive mechanisms of gaze-expression interactions in face processing and social attention". In: *Neuropsychologia* 50.5 (2012), pp. 553–566.

[45] Saul Greenberg and David Marwood. "Real time groupware as a distributed system: concurrency control and its effect on the interface". In: *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. ACM, 1994, pp. 207–217. ISBN: 0-89791-689-1. DOI: 10.1145/192844.193011.

[46] Tovi Grossman, Justin Matejka, and George Fitzmaurice. "Chronicle: capture, exploration, and playback of document workflow histories". In: *Proceedings of the 23nd annual ACM symposium on User interface software and technology*. ACM. 2010, pp. 143–152.

[47] Jens Grubert, Matthias Heinisch, Aaron Quigley, and Dieter Schmalstieg. "Multifi: Multi fidelity interaction with displays on and around the body". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM. 2015, pp. 3933–3942.

[48] Carl Gutwin and Saul Greenberg. "A descriptive framework of workspace awareness for real-time groupware". In: *Computer Supported Cooperative Work (CSCW)* 11.3-4 (2002), pp. 411–446.

[49] Carl Gutwin and Saul Greenberg. "Design for individuals, design for groups: tradeoffs between power and workspace awareness". In: (1998).

[50] Carl Gutwin and Saul Greenberg. "The importance of awareness for team cognition in distributed collaboration". In: *Team Cognition: Process and Performance at the Inter- and Intra-individual Level*. Ed. by E. Salas, S. M. Fiore, and J.A. Cannon-Bowers. 2001.

[51] Carl Gutwin, Oliver Schneider, Robert Xiao, and Stephen Brewster. "Chalk sounds: the effects of dynamic synthesized audio on workspace awareness in distributed groupware". In: *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. ACM. 2011, pp. 85–94.

[52] Carl Gutwin, Gwen Stark, and Saul Greenberg. "Support for workspace awareness in educational groupware". In: *CSCL*. Vol. 95. 1995, pp. 147–156.

[53] Nigel Habershon. "Metaplan (R): achieving two-way communications". In: *Journal of European Industrial Training* 17.7 (1993).

[54] Martin Hachet, Benoit Bossavit, Aurélie Cohé, and Jean-Baptiste de la Rivière. "Toucheo: multitouch and stereo combined in a seamless workspace". In: *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 2011, pp. 587–592. ISBN: 978-1-4503-0716-1. DOI: 10.1145/2047196.2047273.

[55] Shaza Hakim. *The Users Don't Know What They Want, or Do They?—Asking the Right Questions for UX*. URL: https://medium.com/stampede-team/the-users-dont-know-what-they-want-or-do-they-asking-the-right-questions-for-ux-49969607654b. 2017.

[56] Jonna Häkkilä, Olli Koskenranta, Maaret Posti, Leena Ventä-Olkkonen, and Ashley Colley. "Clearing the virtual window: connecting two locations with interactive public displays". In: *Proceedings of the 2nd ACM International Symposium on Pervasive Displays*. ACM. 2013, pp. 85–90.

[57] Chris Harrison, Shilpa Ramamurthy, and Scott E. Hudson. "On-body interaction: armed and dangerous". In: *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*. Kingston, Ontario, Canada: ACM, 2012, pp. 69–76. ISBN: 978-1-4503-1174-8. DOI: 10.1145/2148131.2148148.

[58] Sandra G Hart and Lowell E Staveland. "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research". In: *Advances in psychology*. Vol. 52. Elsevier, 1988, pp. 139–183.

[59] Jörg Hauber, Holger Regenbrecht, Mark Billinghurst, and Andy Cockburn. "Spatiality in videoconferencing: trade-offs between efficiency and social presence". In: *CSCW '06 Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. ACM New York, NY, USA, 2006, pp. 413–422.

[60] Keita Higuch, Ryo Yonetani, and Yoichi Sato. "Can eye help you?: effects of visualizing eye fixations on remote collaboration scenarios for physical tasks". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM. 2016, pp. 5180–5190.

[61] Keita Higuchi, Yinpeng Chen, Philip A. Chou, Zhengyou Zhang, and Zicheng Liu. "Immerseboard: Immersive telepresence experience using a digital whiteboard". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2015. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702160.

[62] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. "Consumed endurance: a metric to quantify arm fatigue of mid-air interactions". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Melbourne, Australia: ACM, 2014, pp. 1063–1072. ISBN: 978-1-4503-2473-1. DOI: 10.1145/2556288.2557130.

[63] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. "Consumed endurance: a metric to quantify arm fatigue of mid-air interactions". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2014, pp. 1063–1072.

[64] H.G. Hoffman. "Physically touching virtual objects using tactile augmentation enhances the realism of virtual environments". In: *IEEE 1998 Virtual Reality Annual International Symposium*. Atlanta, GA, USA, USA: IEEE, 1998. ISBN: 0-8186-8362-7. DOI: 10.1109/VRAIS.1998.658423.

[65] Oliver Hohlfeld, André Pomp, Jó Ágila Bitsch Link, and Dennis Guse. "On the applicability of computer vision based gaze tracking in mobile scenarios". In: *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM. 2015, pp. 427–434.

[66] Chien-Ming Huang, Sean Andrist, Allison Sauppé, and Bilge Mutlu. "Using gaze patterns to predict task intent in collaboration". In: *Frontiers in psychology* 6 (2015), p. 1049.

[67] H Ishiguro. "The uncanny advantage of using androids in social and cognitive science Research". In: (2006).

[68] Hiroshi Ishii and Minoru Kobayashi. "ClearBoard: a seamless medium for shared drawing and conversation with eye contact". In: *CHI '92 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1992, pp. 525–532. ISBN: 0-89791-513-5. DOI: 10.1145/142750.142977.

[69] Shahram Izadi, Ankur Agarwal, Antonio Criminisi, John Winn, Andrew Blake, and Andrew Fitzgibbon. "C-Slate: a multi-touch and object recognition system for remote collaboration using horizontal surfaces". In: *Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'07)*. IEEE. 2007, pp. 3–10.

[70] Mikkel R. Jakobsen, Yvonne Jansen, Sebastian Boring, and Kasper Hornbæk. "Should I Stay or Should I Go? Selecting Between Touch and Mid-Air Gestures for Large-Display Interaction". In: *IFIP Conference on Human-Computer Interaction*. Springer, 2015. ISBN: 978-3-319-22697-2. DOI: 10.1007/978-3-319-22698-9_31.

[71] Kristiina Jokinen. "Gaze and gesture activity in communication". In: *International Conference on Universal Access in Human-Computer Interaction*. Springer. 2009, pp. 537–546.

[72] Annabel Bhamani Kajornboon. "Using interviews as research instruments". In: *E-journal for Research Teachers* 2.1 (2005), pp. 1–9.

[73] Adam Kendon. "Some functions of gaze-direction in social interaction". In: *Acta Psychologica* 26 (1967).

[74] Thomas Kennel, Andreas Kunz, and Stephan Müller. "Innoplan—An adaptation of the Metaplan technique for a novel computer supported method of teamwork". In: *Design research. Theories, methodologies, and product modelling* 28 (2001), pp. 387–394.

[75] Kibum Kim, John Bolton, Audrey Girouard, Jeremy Cooperstock, and Roel Vertegaal. "TeleHuman: effects of 3d perspective on gaze and pose estimation with a life-size cylindrical telepresence pod". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2012, pp. 2531–2540.

[76] Seungwon Kim, Gun A Lee, Sangtae Ha, Nobuchika Sakata, and Mark Billinghurst. "Automatically freezing live video for annotation during remote collaboration". In: *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM. 2015, pp. 1669–1674.

[77] David Kirk, Andy Crabtree, and Tom Rodden. "Ways of the hands". In: *ECSCW 2005*. Springer. 2005, pp. 1–21.

[78] 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*. ACM. 2006, pp. 1191–1200.

[79] Scott Klemmer. *Introduction to Human-Computer Interaction Design*. URL: `https://d.ucsd.edu/class/intro-hci/2014/index.html`. 2015.

[80] Scott R Klemmer, Michael Thomsen, Ethan Phelps-Goodman, Robert Lee, and James A Landay. "Where do web sites come from?: capturing and interacting with design history". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2002, pp. 1–8.

[81] Benjamin Koehne, Patrick C Shih, and Judith S Olson. "Remote and alone: coping with being the remote member on the team". In: *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. ACM. 2012, pp. 1257–1266.

[82] Martin Kuechler and Andreas Kunz. "Holoport-a device for simultaneous video and data conferencing featuring gaze awareness". In: *Virtual Reality Conference 2006*. IEEE, 2006. ISBN: 1-4244-0224-7. DOI: `10.1109/VR.2006.71`.

[83] Martin Kuechler and Andreas M Kunz. "Collaboard: a remote collaboration groupware device featuring an embodiment-enriched shared workspace". In: *Proceedings of the 16th ACM international conference on Supporting group work*. ACM. 2010, pp. 211–214.

[84] Andreas Kunz, Thomas Nescher, and Martin Kuchler. "Collaboard: a novel interactive electronic whiteboard for remote collaboration with people on content". In: *2010 International Conference on Cyberworlds*. IEEE, 2010. ISBN: 978-1-4244-8301-3. DOI: `10.1109/CW.2010.17`.

[85] Ricardo Langner, Tom Horak, and Raimund Dachselt. "VisTiles: coordinating and combining co-located mobile devices for visual data exploration". In: *IEEE Transactions on Visualization and Computer Graphics* 24.2 (Aug. 2017), pp. 626–636.

[86] Ricardo Langner, Ulrike Kister, and Raimund Dachselt. "Multiple coordinated views at large displays for multiple users: Empirical findings on user behavior, Movements, and Distances". In: *IEEE Transactions on Visualization and Computer Graphics* (Aug. 2018). DOI: `10.1109/TVCG.2018.2865235`.

[87] Khanh-Duy Le, Kening Zhu, and Morten Fjeld. "Mirrortablet: exploring a low-cost mobile system for capturing unmediated hand gestures in remote collaboration". In: *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia*. ACM. 2017, pp. 79–89.

[88] Min Kyung Lee and Leila Takayama. "Now, I have a body: Uses and social norms for mobile remote presence in the workplace". In: *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM. 2011, pp. 33–42.

[89] Daniel Leithinger, Sean Follmer, Alex Olwal, and Hiroshi Ishii. "Physical telepresence: shape capture and display for embodied, computer-mediated remote collaboration". In: *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM, 2014, pp. 461–470. ISBN: 978-1-4503-3069-5. DOI: 10.1145/2642918.2647377.

[90] Jiannan Li, Saul Greenberg, Ehud Sharlin, and Joaquim Jorge. "Interactive two-sided transparent displays: designing for collaboration". In: *Proceedings of the 2014 conference on Designing interactive systems*. ACM. 2014, pp. 395–404.

[91] Yingjing Jane Li. "Designing collaborative workspaces for particular complex work settings". PhD thesis. 2016.

[92] Rensis Likert. "A technique for the measurement of attitudes." In: *Archives of psychology* (1932).

[93] Can Liu, Olivier Chapuis, Michel Beaudouin-Lafon, and Eric Lecolinet. "CoReach: Cooperative gestures for data manipulation on wall-sized displays". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM. 2017, pp. 6730–6741.

[94] Andrés Lucero, Matt Jones, Tero Jokela, and Simon Robinson. "Mobile collocated interactions: taking an offline break together". In: *interactions* 20.2 (2013), pp. 26–32.

[95] Andrés Lucero, Jaakko Keränen, and Hannu Korhonen. "Collaborative use of mobile phones for brainstorming". In: *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*. ACM. 2010, pp. 337–340.

[96] Antoinette L Lynch, Uday S Murthy, and Terry J Engle. "Fraud brainstorming using computer-mediated communication: The effects of brainstorming technique and facilitation". In: *The Accounting Review* 84.4 (2009), pp. 1209–1232.

[97] Nicolai Marquardt, Frederik Brudy, Can Liu, Ben Bengler, and Christian Holz. "SurfaceConstellations: a modular hardware platform for ad-hoc reconfigurable cross-device workspaces". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. 2018, p. 354.

[98] Nicolai Marquardt, Ken Hinckley, and Saul Greenberg. "Cross-device interaction via micro-mobility and f-formations". In: *Proceedings of the 25th annual ACM symposium on User interface software and technology*. ACM. 2012, pp. 13–22.

[99] Lisa Min. *It's True: Users Don't Know What They Want*. URL: https://www.akendi.com/blog/its-true-users-dont-know-what-they-want/. 2016.

[100] Peter Mohr, Markus Tatzgern, Tobias Langlotz, Andreas Lang, Dieter Schmalstieg, and Denis Kalkofen. "TrackCap: Enabling smartphones for 3D interaction on mobile head-mounted displays". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM. 2019, p. 585.

[101] Jörg Müller, Dieter Eberle, and Konrad Tollmar. "Communiplay: a field study of a public display mediaspace". In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM. 2014, pp. 1415–1424.

[102] Thomas Neumayr, Hans-Christian Jetter, Mirjam Augstein, Judith Friedl, and Thomas Luger. "Domino: A descriptive framework for hybrid collaboration and coupling styles in partially distributed teams". In: *Proceedings of the ACM on Human-Computer Interaction* 2.CSCW (2018), p. 128.

[103] David Nguyen and John Canny. "MultiView: spatially faithful group video conferencing". In: *CHI '05 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM New York, NY, USA, 2005, pp. 799–808.

[104] Jakob Nielsen and Rolf Molich. "Heuristic evaluation of user interfaces". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 1990, pp. 249–256.

[105] Donald A. Norman. "Things that make us smart". In: 1993.

[106] David G Novick, Brian Hansen, and Karen Ward. "Coordinating turn-taking with gaze". In: *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP'96*. Vol. 3. IEEE. 1996, pp. 1888–1891.

[107] Anne K Offner, Thomas J Kramer, and Joel P Winter. "The effects of facilitation, recording, and pauses on group brainstorming". In: *Small group research* 27.2 (1996), pp. 283–298.

[108] Gary M. Olson and Judith S. Olson. "Distance matters". In: *Human-Computer Interaction* 15.2-3 (Dec. 2009), pp. 139–178.

[109] Kazuhiro Otsuka. "MMSpace: Kinetically-augmented telepresence for small group-to-group conversations". In: *Proceedings of 2016 IEEE Virtual Reality (VR)*. IEEE, 2016.

[110] Mai Otsuki, Taiki Kawano, Keita Maruyama, Hideaki Kuzuoka, and Yusuke Suzuki. "ThirdEye: Simple add-on display to represent remote participant's gaze direction in video communication". In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2017, pp. 5307–5312. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025681.

[111] Mai Otsuki, Keita Maruyama, Hideaki Kuzuoka, and Yusuke Suzuki. "Effects of enhanced gaze presentation on gaze leading in remote collaborative physical tasks". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. 2018, p. 368.

[112] Jeni Paay, Dimitrios Raptis, Jesper Kjeldskov, Mikael B Skov, Eric V Ruder, and Bjarke M Lauridsen. “Investigating cross-device interaction between a handheld device and a large display”. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM. 2017, pp. 6608–6619.

[113] Gian Pangaro, Dan Maynes-Aminzade, and Hiroshi Ishii. “The actuated workbench: computer-controlled actuation in tabletop tangible interfaces”. In: *Proceedings of the 15th annual ACM symposium on User interface software and technology*. ACM, 2002, pp. 461–470. ISBN: 1-58113-488-6. DOI: 10.1145/571985.572011.

[114] Sébastien Piérard, Vincent Pierlot, Antoine Lejeune, and Marc Van Droogenbroeck. “I-see-3d! an interactive and immersive system that dynamically adapts 2d projections to the location of a user’s eyes”. In: *2012 International Conference on 3D Imaging (IC3D)*. IEEE. 2012, pp. 1–8.

[115] Andrea C Pierno, Cristina Becchio, Matthew B Wall, Andrew T Smith, Luca Turella, and Umberto Castiello. “When gaze turns into grasp”. In: *Journal of Cognitive Neuroscience* 18.12 (2006), pp. 2130–2137.

[116] David Pinelle, Miguel Nacenta, Carl Gutwin, and Tadeusz Stach. “The effects of co-present embodiments on awareness and collaboration in tabletop groupware”. In: *GI ’08 Proceedings of Graphics Interface 2008*. ACM, 2008, pp. 1–8. ISBN: 978-1-56881-423-0.

[117] Thammathip Piumsomboon, Gun A Lee, Jonathon D Hart, Barrett Ens, Robert W Lindeman, Bruce H Thomas, and Mark Billinghurst. “Mini-me: An adaptive avatar for mixed reality remote collaboration”. In: *Proceedings of the 2018 CHI conference on human factors in computing systems*. ACM. 2018, p. 46.

[118] Martin Porcheron, Andrés Lucero, and Joel E Fischer. “Co-curator: designing for mobile ideation in groups”. In: *Proceedings of the 20th International Academic Mindtrek Conference*. ACM. 2016, pp. 226–234.

[119] Jenny Preece, Yvonne Rogers, and Helen Sharp. *Interaction design: beyond human-computer interaction*. John Wiley & Sons, 2015.

[120] Arnaud Prouzeau, Anastasia Bezerianos, and Olivier Chapuis. “Awareness Techniques to Aid Transitions between Personal and Shared Workspaces in Multi-Display Environments”. In: *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces*. ACM. 2018, pp. 291–304.

[121] Arnaud Prouzeau, Anastasia Bezerianos, and Olivier Chapuis. “Towards road traffic management with forecasting on wall displays”. In: *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*. ACM. 2016, pp. 119–128.

[122] Rebecca Randell, Thilina Ambepitiya, Claudia Mello-Thoms, Roy A Ruddle, David Brettle, Rhys G Thomas, and Darren Treanor. "Effect of display resolution on time to diagnosis with virtual pathology slides in a systematic search task". In: *Journal of digital imaging* 28.1 (2015), pp. 68–76.

[123] Khairi Reda, Andrew E Johnson, Michael E Papka, and Jason Leigh. "Effects of display size and resolution on user behavior and insight acquisition in visual exploration". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM. 2015, pp. 2759–2768.

[124] Holger Regenbrecht, Michael Haller, Joerg Hauber, and Mark Billinghurst. "Carpeno: interfacing remote collaborative virtual environments with table-top interaction". In: *Virtual Reality* 10.2 (2006), pp. 95–107.

[125] Peter Robinson and Philip Tuddenham. "Distributed Tabletops: Supporting remote and mixed-presence tabletop collaboration". In: *Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP'07)*. IEEE, 2007. ISBN: 978-0-7695-2013-1. DOI: 10.1109/TABLETOP.2007.15.

[126] Yong Rui, Anoop Gupta, and Jonathan J Cadiz. "Viewing meeting captured by an omni-directional camera". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2001, pp. 450–457.

[127] Tony Salvador, Jean Scholtz, and James Larson. "The Denver model for groupware design". In: *ACM SIGCHI Bulletin* 28.1 (Jan. 1996), pp. 52–58.

[128] Ugo Braga Sangiorgi, François Beuvens, and Jean Vanderdonckt. "User interface design by collaborative sketching". In: *Proceedings of the Designing Interactive Systems Conference*. ACM. 2012, pp. 378–387.

[129] Kjeld Schmidt. "The Problem with 'Awareness'". In: *Cooperative Work and Coordinative Practices*. Springer, 2002, pp. 157–166.

[130] Jean Scholtz. "Usability evaluation". In: *National Institute of Standards and Technology* 1 (2004).

[131] L. Segal. "Designing team workstations: The choreography of teamwork". In: *Local Applications of the Ecological Approach to Human-Machine Systems*. Ed. by P. Hancock, J. Flach, J. Caird, and K. Vicente. 1995, pp. 392–415.

[132] Manuel César Bessa Seixas, Jorge CS Cardoso, and Maria Teresa Galvão Dias. "The Leap Motion movement for 2D pointing tasks: Characterisation and comparison to other devices". In: *2015 International Conference on Pervasive and Embedded Computing and Communication Systems (PECCS)*. IEEE. 2015, pp. 15–24.

[133] Jun'ichiro Seyama and Ruth S. Nagayama. "The uncanny valley: Effect of realism on the impression of artificial human faces". In: *PRESENCE: Teleoperators and Virtual Environments* 16.4 (2007), pp. 337–351. DOI: 10.1162/pres.16.4.337.

[134] Jun'ichiro Seyama and Ruth S Nagayama. "The uncanny valley: Effect of realism on the impression of artificial human faces". In: *Presence: Teleoperators and virtual environments* 16.4 (2007), pp. 337–351.

[135] Rajinder S. Sodhi, Brett R. Jones, David Forsyth, Brian P. Bailey, and Giuliano Maciocci. "BeThere: 3D mobile collaboration with spatial input". In: *CHI '13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2013, pp. 179–188. ISBN: 978-1-4503-1899-0. DOI: 10.1145/2470654.2470679.

[136] Martin Spindler, Wolfgang Büschel, and Raimund Dachselt. "Use your head: tangible windows for 3D information spaces in a tabletop environment". In: *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces*. ACM. 2012, pp. 245–254.

[137] Susan Leigh Star and Karen Ruhleder. "Steps towards an ecology of infrastructure: complex problems in design and access for large-scale collaborative systems". In: *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. Chapel Hill, North Carolina, USA: ACM, 1994, pp. 253–264. ISBN: 0-89791-689-1. DOI: 10.1145/192844.193021.

[138] Nawel Takouachet, Jérémy Legardeur, and Iban Lizarralde. "The role of the facilitator during digital creative sessions". In: *Proceedings of the 2014 Ergonomie et Informatique Avancée Conference-Design, Ergonomie et IHM: quelle articulation pour la co-conception de l'interaction*. New York, NY, USA: ACM, 2014, pp. 20–23.

[139] John C. Tang and Scott Minneman. "VideoWhiteboard: video shadows to support remote collaboration". In: *CHI '91 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1991, pp. 315–322. ISBN: 0-89791-383-3. DOI: 10.1145/108844.108932.

[140] John C. Tang and Scott L. Minneman. "VideoDraw: a video interface for collaborative drawing". In: *CHI '90 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1990, pp. 313–320. ISBN: 0-201-50932-6. DOI: 10.1145/97243.97302.

[141] Hiroaki Tobita, Shigeaki Maruyama, and Takuya Kuzi. "Floating avatar: telepresence system using blimps for communication and entertainment". In: *CHI'11 Extended Abstracts on Human Factors in Computing Systems*. ACM. 2011, pp. 541–550.

[142] Jayson Turner, Jason Alexander, Andreas Bulling, and Hans Gellersen. "Gaze+RST: integrating gaze and multitouch for remote rotate-scale-translate tasks". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM. 2015, pp. 4179–4188.

[143] Gina Venolia, John Tang, Ruy Cervantes, Sara Bly, George Robertson, Bongshin Lee, and Kori Inkpen. "Embodied social proxy: mediating interpersonal connection in hub-and-satellite teams". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2010, pp. 1049–1058.

[144] Roel Vertegaal. "The GAZE groupware system: mediating joint attention in multiparty communication and collaboration". In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM. 1999, pp. 294–301.

[145] Roel Vertegaal, Robert Slagter, Gerrit Van der Veer, and Anton Nijholt. "Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2001, pp. 301–308.

[146] Matthew O Ward, Georges Grinstein, and Daniel Keim. *Interactive data visualization: foundations, techniques, and applications*. AK Peters/CRC Press, 2015.

[147] Robert W Weisberg. "I2 Creativity and Knowledge: A challenge t0 the0ries". In: *Handbook of creativity* 226 (1999).

[148] Eric Whitmire, Hrvoje Benko, Christian Holz, Eyal Ofek, and Mike Sinclair. "Haptic Revolver: Touch, shear, texture, and shape rendering on a reconfigurable virtual reality controller". In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. 2018, p. 86.

[149] Chauncey Wilson. *Interview techniques for UX practitioners: A user-centered design method*. Newnes, 2013.

[150] Erroll Wood and Andreas Bulling. "Eyetab: Model-based gaze estimation on unmodified tablet computers". In: *Proceedings of the Symposium on Eye Tracking Research and Applications*. ACM. 2014, pp. 207–210.

[151] Erroll Wood, Jonathan Taylor, John Fogarty, Andrew Fitzgibbon, and Jamie Shotton. "ShadowHands: High-fidelity remote hand gesture visualization using a hand tracker". In: *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*. ACM, 2016, pp. 77–84. ISBN: 978-1-4503-4248-3. DOI: 10.1145/2992154.2992169.

[152] Paweł Wozniak, Nitesh Goyal, Przemysław Kucharski, Lars Lischke, Sven Mayer, and Morten Fjeld. "RAMPARTS: Supporting sensemaking with spatially-aware mobile interactions". In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM. 2016, pp. 2447–2460.

[153] Paweł Woźniak, Lars Lischke, Benjamin Schmidt, Shengdong Zhao, and Morten Fjeld. "Thaddeus: a dual device interaction space for exploring information visualisation". In: *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*. ACM. 2014, pp. 41–50.

[154] Peter Wright and John McCarthy. "Empathy and experience in HCI". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2008, pp. 637–646.

[155] Dustin Wüest, Norbert Seyff, and Martin Glinz. “Sketching and notation creation with FlexiSketch Team: Evaluating a new means for collaborative requirements elicitation”. In: *2015 IEEE 23rd International Requirements Engineering Conference (RE)*. IEEE. 2015, pp. 186–195.

[156] Robert Xiao, Julia Schwarz, Nick Throm, Andrew D. Wilson, and Hrvoje Benko. “MRTouch: Adding touch input to head-mounted mixed reality”. In: *IEEE Transactions on visualization and computer graphics* 24.4 (Apr. 2018), pp. 1653–1660.

[157] Bin Xu, Jason Ellis, and Thomas Erickson. “Attention from afar: Simulating the gazes of remote participants in hybrid meetings”. In: *Proceedings of the 2017 Conference on Designing Interactive Systems*. ACM. 2017, pp. 101–113.

[158] Naomi Yamashita, Katsuhiko Kaji, Hideaki Kuzuoka, and Keiji Hirata. “Improving visibility of remote gestures in distributed tabletop collaboration”. In: *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. ACM. 2011, pp. 95–104.

[159] Naomi Yamashita, Hideaki Kuzuoka, Keiji Hirata, Shigemi Aoyagi, and Yoshinari Shirai. “Supporting fluid tabletop collaboration across distances”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2011, pp. 2827–2836.

[160] Jakob Zillner, Christoph Rhemann, Shahram Izadi, and Michael Haller. “3D-board: a whole-body remote collaborative whiteboard”. In: *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM. 2014, pp. 471–479.