UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

JERÔNIMO GUSTAVO GRANDI

Collaborative 3D Interactions and Their Application on Virtual, Augmented and Mixed Reality Interfaces

Thesis presented in partial fulfillment of the requirements for the degree of Doctor of Computer Science

Advisor: Prof. Dr. Anderson Maciel

Porto Alegre
December 2018

CIP — CATALOGING-IN-PUBLICATION

Grandi, Jerônimo Gustavo

Collaborative 3D Interactions and Their Application on Virtual, Augmented and Mixed Reality Interfaces / Jerônimo Gustavo Grandi. – Porto Alegre: PPGC da UFRGS, 2018.

123 f.: il.

Thesis (Ph.D.) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR–RS, 2018. Advisor: Anderson Maciel.

1. 3D User Interface. 2. Collaborative Object Manipulation. 3. Augmented Reality. 4. Virtual Reality. 5. Asymmetric Interaction. I. Maciel, Anderson. II. Título.

ACKNOWLEDGMENTS

I wish to address a special thank to my Advisor Prof. Anderson Maciel, Ph.D. to encouraging my research and allowing me to pursue my objectives. I also thank Prof. Luciana Nedel, Ph.D. for the collaboration in all phases of my project and all valuable knowledge shared.

I thank the Federal University of Rio Grande do Sul (UFRGS) and the Institute of Informatics (INF) that provided the infrastructure and funding for the execution of my research. Additionally, I thank CAPES and CNPq-Brazil for the financial support to the provision of post-graduate scholarship. I also acknowledge FAPERGS (project PqG 2283-2551/14-8 and PqG 17/2551-0001), CNPq-Brazil (project PQ-2 311353/2017-7) and Petrobras (project PD ANP 2015/00068-0).

I want to thank all my lab colleagues and friends. It was an awesome experience to do research at the Graphics, Visualization and Interaction Group, to meet people sharing the same passion and for all fun moments. A special thank to Victor A. J. Oliveira for their precious feedbacks on my work, help on the experiment design and for the great discussions; to Henrique G. Debarba for the valuable collaboration on all my papers, helping me in writing and statistical analysis; to Iago Berndt for the commitment and collaboration in the contests; to Sammer M. C. G. Dias for the support, love and patience.

To my family, thanks for all the assistance and love. You are the main reason for my success. I owe you a lot, and my gratitude is beyond words, this thesis is dedicated to them.

ABSTRACT

Spatial object manipulation is one of the most common tasks in 3D environments. It is the action of changing position, rotation and scale properties of virtual elements. The intangibility of the virtual requires the use of 3D user interfaces for robust interactions. Manipulations in 3D involve complex coordination of virtual transformations and demand a highly cognitive and physical effort. One alternative towards efficient 3D manipulations is to share the work to solve the task collaboratively. A collaborative environment provides more than one individual perspectives of the scene, which may also include shared and personal viewpoints. However, teamwork involves considerable negotiations and, as team members vary, team strategies and task-fulfillment processes also change. Modeling such human interactions raises novel concepts and interface design approaches when compared to those typically grounded in a single-user scenario.

Our goal is to investigate collaboration in group work, identify the aspects of effective cooperative interactions and design 3D user interactions that enhance the team capability to solve complex manipulation tasks in virtual environments. We explore the use of handheld devices and virtual reality apparatus as input devices in our designs. Smartphones are ubiquitous in modern life which makes them useful for collaboration, and virtual reality headsets are becoming the standard for immersive virtual reality. First, we designed an interaction model to support simultaneous 3D manipulations in virtual and augmented reality scenarios and evaluated group performance and behavior in symmetric collaboration, where the team has the same interaction and visualization interface. After, we investigate asymmetric group interactions, integrating our symmetric collaborative solutions. This research contributes to the definition of a model for the design of collaborative 3D user interfaces flexible to work in multiple immersion scenarios. Furthermore, we present a series of studies that investigate the application of collaborative 3D interactions on Virtual, Augmented and Mixed Reality scenarios.

Keywords: 3D User Interface. Collaborative Object Manipulation. Augmented Reality. Virtual Reality. Asymmetric Interaction.

Interações 3D Colaborativas e Sua Aplicação em Interfaces de Realidade Virtual, Aumentada e Mista

RESUMO

A manipulação espacial de objetos é uma das tarefas mais comuns em ambientes 3D. É a ação de alterar as propriedades de posição, rotação e escala dos elementos virtuais. A intangibilidade do virtual requer o uso de interfaces de usuário 3D para interações robustas. Manipulações em 3D envolvem a coordenação complexa de transformações virtuais e exigem um alto esforço cognitivo e físico. Uma alternativa para manipulações 3D eficientes é compartilhar o trabalho para resolver a tarefa de forma colaborativa. Um ambiente colaborativo fornece uma perspectiva individual da cena, que também pode incluir pontos de vista compartilhados e pessoais. No entanto, o trabalho em equipe envolve negociações consideráveis e, à medida que os membros da equipe variam, as estratégias de equipe e os processos de realização de tarefas também mudam. A modelagem de tais interações humanas gera novos conceitos e abordagens de design de interface quando comparados àqueles normalmente baseados em um cenário de um único usuário. Nosso objetivo é investigar a colaboração no trabalho em grupo, identificar os aspectos de interações cooperativas efetivas e projetar interações de usuários 3D que aprimorem a capacidade da equipe de resolver tarefas complexas de manipulação em ambientes virtuais. Exploramos o uso de dispositivos portáteis e aparelhos de realidade virtual como interfaces de entrada em nossos projetos. Os smartphones são onipresentes na vida moderna, o que os torna úteis para a colaboração, e os capacetes de realidade virtual estão se tornando o padrão para a realidade virtual imersiva. Primeiro, projetamos um modelo de interação para suportar manipulações 3D simultâneas em cenários de realidade virtual e aumentada e avaliamos o desempenho e o comportamento do grupo em colaboração simétrica, em que a equipe tem a mesma interação e interface de visualização. Depois, investigamos interações assimétricas em grupos, integrando nossas soluções colaborativas simétricas. Esta pesquisa contribui para a definição de um modelo para o design de interfaces de usuário 3D colaborativas, flexível para trabalhar em múltiplos cenários de imersão. Além disso, apresentamos o uma série de estudos que investigam a aplicação de interações 3D colaborativas em cenários de Realidade Virtual, Aumentada e Mista.

Palavras-chave: Interfaces 3D de Usuário, Manipulação Colaborativa de Objetos 3D, Realidade Aumentada, Realidade Virtual, Interação Assimétrica.

LIST OF ABBREVIATIONS AND ACRONYMS

3D	Three dimensional
AR	Augmented reality
CVE	Collaborative virtual environment
CSCW	Computer Supported Collaborative Work
DOF	Degree of freedom
HMD	Head-Mounted Display
NASA TLX	NASA Task Load Index
PoV	Point-of-View
RST	Rotate-scale-transform
SEQ	Single Easy Question
SoC	Sum-of-Contributions
SUS	System Usability Scale
TUI	Tangible User Interface
UI	User Interface
VE	Virtual Environment
VR	Virtual reality

LIST OF FIGURES

CONTENTS

1 INTRODUCTION

Humans cooperate on a broad range of activities. They can, for instance, organize themselves into complex labor divisions to accomplish tasks that alone would be difficult. Cooperation can enhance productivity and deliver faster results (LAUGHLIN, 2011). Nowadays, a considerable amount of work is supported by computers, and naturally, we need to evolve the tools to support collaboration in such digital spaces. Collaborative Virtual Environments (CVEs) are shared spaces designed to support interactions between users and objects in Virtual Reality (VR) (CHURCHILL; SNOWDON, 1998). Due to the unlimited nature of the virtual, besides only simulate the aspects of the physical world, it is possible to extrapolate the bounds of the real to enhance user's abilities and perception during the virtual reality experience. CVEs add the possibility to integrate mixed realities where users can co-exist in an asymmetric experience.

Virtual 3D manipulations are one of the canonical tasks in VEs (Bowman A. D; Kruijff, 2005). It is the action of changing position, rotation and scale properties of virtual elements. Virtual objects are intangible in 3D VEs, and the interaction with them is only achieved through full-body tracking or mediated by a user interface device. Creating effective interfaces for 3D virtual systems is challenging (ZHAI, 1995; STUERZLINGER; WINGRAVE, 2011). The capacity of interaction depends on the mapping of user inputs into the correct action. Moreover, the visualization with projections, stereoscopic displays and HMD offer different levels of immersion and affect the user experience when interacting with virtual objects. In 3D manipulation tasks, it is hard to be precise when working from a single perspective. The accomplishment of simple 3D manipulations, such as the precise positioning of an object, requires the continuous movement of the user to check for occluded parts. This is tedious and time-consuming, especially when the objects' shape is complex. While these manipulations are mandatory on any 3D interactive virtual environment, performing them alone demands a high cognitive effort to coordinate the possible transformations (VEIT; CAPOBIANCO; BECHMANN, 2009). Many 3D User Interfaces (3DUI) have been proposed to allow natural, precise and fast interaction with virtual environments (Bowman A. D; Kruijff, 2005). Techniques are typically designed for specific purpose applications. Some of them have become relatively common due to the easy access to the market, such as gaming console interfaces. However, as of today, research in the 3D object manipulation field mainly focus on single-user interaction.

Since it is interesting that multiple people share the same virtual space and cooperate in there, an interface capable of handling inputs of many users for cooperative work is desirable. One of the main advantages of having a collaborative interface is that it provides users with an individual action perspective of the same scene, which may also include shared and individual viewpoints. Teamwork involves considerable negotiations and, as team members vary, team strategies and task-fulfillment processes also change. Modeling such interactions raises novel collaborative concepts compared to those typically grounded in a single-user scenario (ROBINSON, 1993).

1.1 Objectives and Approach

Our goal is to explore collaborative work for 3D manipulation tasks of virtual objects. A major research problem of collaborative activity is to manage shared access to the same object for simultaneous manipulations (MARGERY; ARNALDI; PLOUZEAU, 1999). For instance, in the *Piano Mover's* problem (LENGYEL et al., 1990), where a piano has to be moved through a corridor while avoiding collisions, two or more users need to negotiate actions, control forces and coordinate movements to successfully carry the piano to the final destination. In this thesis, we research and propose efficient simultaneous manipulations and parallel work. We conduct user-guided studies to evaluate collaborative interfaces. In those assessments, we address open research problems that are fundamental for seamless cooperative interactions: tasks division and group organization, group sizes for optimal productivity, individual and group awareness, and individual and group responsibility for cooperative interactions in 3D virtual spaces.

We applied our *Sum-of-Contributions* model in the design of collaborative 3DUIs using recent and affordable hardware: Mobile Devices and Head-Mounted Displays (HMDs). Mobile devices are not designed for interaction in VEs. However, they encapsulate several characteristics that make them efficient 3D interaction tools, such as, processing power, inertial sensors, wireless connection and touchscreens. Moreover, smartphones are ubiquitous in modern life and are likely to be the most widely available input device, which is desirable when designing virtual environments that work across different platforms. Additionally, we use mobile devices as an extension of our body (BALLAGAS et al., 2006), and we can assume that users of all ages and backgrounds have developed familiarity with the mobile devices input interfaces, which could result in a relatively short learning curve when their features are explored in the 3DUI context. On the other hand, HMDs are conceived for 3D immersive experiences.

In the end, we provide collaborative solutions for the main display platforms, involving augmented and virtual reality. Moreover, individual solutions are compatible with each other. In this way, they can be mixed in various configurations to adapt to the task needs. An unlimited number of users with different interfaces could work together. For instance, some collaborators could work immersed with HMD's while others could visualize the task on a screen or with augmented reality devices. This research also contributes to the definition of baselines for the design of collaborative interactions on VEs.

1.2 Thesis Organization Overview

We organized this thesis as follows: Chapter 2 presents the background and a literature review of 3D interactions and collaborative manipulations in virtual environments. Based on the knowledge acquired with the literature review, we designed a model explore collaborative aspects in four main areas: 3D environments (Chapter 3), AR environments (Chapter 4), immersive VR environments (Chapter 5) and mixed asymmetric environments (Chapter 6). We first restricted our designs with symmetric interactions (first three areas), where all users have the same interface and visualizations available. Thus, we can isolate each approach and assess their strong and weak points individually. After in Chapter 6, we integrated our designs in a mixed asymmetric environment, we evaluate the group performance in an asymmetric approach built upon the already proposed techniques. Finally, Chapter 7 presents a summary of the contributions, design considerations and perspectives of this work.

2 BACKGROUND AND LITERATURE REVIEW

We divided this chapter into three main sections. The first discusses the main approaches for 3D object manipulations using touch surfaces and 6 DOF controllers.Section 2.2 approaches the main concepts of asymmetric interactions in VEs. Then, in Section 2.3, we compile works that apply collaborative aspects for 3D object manipulations and focus on co-manipulations. Finally, in Section 2.4, we introduce and discuss possible contributions in the field.

2.1 3D Manipulation in Virtual Environments

In Virtual Environments (VEs), interaction techniques provide the capability to interact with the virtual elements. The quality of interaction depends on the mapping of user inputs into virtual actions and on the manipulation task which they are applied. Bowman A. D; Kruijff (2005) classifies 3D interaction techniques into four categories: navigation, selection, manipulation and application control. In this thesis, we focus on the manipulation aspects of virtual interactions.

The *isomorphic* approach allow 1:1 correspondence between hand motions in the physical and virtual. While it is generally accepted to be the most natural interaction mapping, it is mostly constrained by the limitations of humans. One advantage of virtual environments is the possibility to extrapolate the bounds of the real to enhance user's abilities. In this direction, *nonisomorphic* approach empowers the user with *magic* tools that deviates significantly from strict realism, while maintaining usability and performance (BOWMAN; HODGES, 1997).

In the following sections, we explore the use of handheld devices and VR controllers for *nonisomorphic* interactions in virtual environments.

2.1.1 Handheld-based Interactions

The use of smartphones as 3DUI for an external computing environment has already been explored to some extent. Here, we focus on works that use touch gestures (Sec. 2.1.1.1), inertial sensors (Sec. 2.1.1.2), around the device manipulations (Sec. 2.1.1.3) and device movements (Sec. 2.1.1.4) as the input for 3D object manipulations.

In the following sections, we address the main techniques to interface user input to 3D transformations in virtual environments. These approaches can be integrated, for instance, surface metaphors are present in all other categories, but we distinct them by their primary contribution.

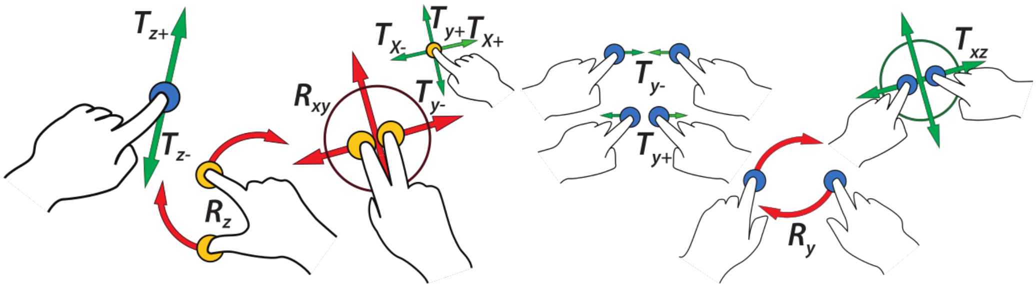
2.1.1.1 Interactions in Touch Surface

Touch gestures are a well-established input for object manipulations. Touchscreens are available in various surface sizes and hardware apparatus, from tabletops to handheld devices. The fast adoption of mobile devices made the manipulations using touch to became natural in people's life. Several designs and usability studies support intuitive interaction with touch. Here, we present works that focus on exploring touch gestures for 3D manipulations of virtual objects. These manipulations can be classified as *direct* or *indirect* depending on how the input happens and how it affects the manipulated object.

Hancock et al. (2006) were the pioneers to address multi-touch for 3D interactions. They proposed direct touch manipulations with one, two or three fingers in their study of *shallow depth* (HANCOCK; CARPENDALE; COCKBURN, 2007). Their results indicate that users are faster, most accurate and prefer the three-touch technique. This work was improved in *Sticky Tools* (HANCOCK; Ten Cate; CARPENDALE, 2009), where the authors modified the third touch to perform rotations around the axis defined by the two other fingers and added physics for force-based effects such as collisions and gravity. Reisman, Davidson and Han (2009) proposed an adaptation for the Rotate-Scale-Transform (RST) 2D metaphor for 3D manipulation with a constraint solver. Their approach preserves the original contact point on an object after each manipulation. This feature facilitates the prediction of how the fingers affect the object's movements.

All techniques described above fall in the "integrated" category where the degrees of freedom (DOF) can be manipulated simultaneously. The separable approach constrains the transformation a user can perform at a time. It has been proposed as an alternative to enhance precision in controlling an integrated manipulation. Notably, Martinet, Casiez and Grisoni (2012) proposed DS3 (Depth-Separated Screen Space), a 3D manipulation technique based on a total separation of the position and orientation control (see Figure 2.1a). The authors used the number of fingers in contact with the screen to determine the transformation, for instance, one finger to control translations and two fingers to control rotations. DS3 also combines the Z-technique (MARTINET; CASIEZ; GRISONI, 2010) to control the position of the Z axis and the constraint solver described by Reisman, Davidson

Figure 2.1: 2D Multi-Touch techniques. (a) direct manipulation, (b) indirect manipulation.



(a) DS3 (b) Triangle Cursor

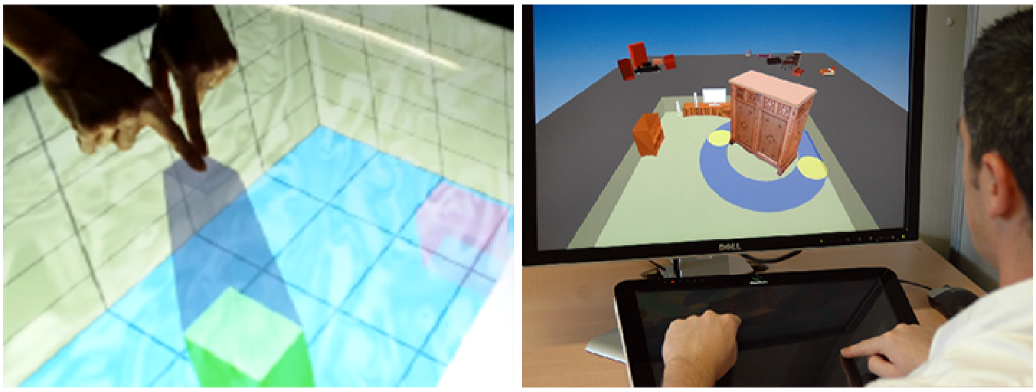
Source: Simeone (2016)

and Han (2009) to control the orientation with an indirect finger.

Display size limits the direct touch techniques when the physical inputs need to cover a large area. Also, many users interacting with a collaborative system can obstruct each other's view and limit the interaction area. Indirect techniques, on the other hand, asymmetrically maps the input-output, overcoming the collaboration problems mentioned earlier. Indirect techniques require a proxy, such as a cursor, a UI element or an external surface, to interact with an object (FORLINES et al., 2007). Indirect approaches can be differentiated between those which the input is on the same surface and those who use an external device to express the input. The use of the same surface for indirect input has been explored by the Balloon technique, that uses the metaphor of a helium balloon controlled by touches in a screen (BENKO; FEINER, 2007), the triangle cursor (STROTHOFF; VALKOV; HINRICHS, 2011) (Figure 2.1b), designed for stereoscopic displays and Void Shadows (GIESLER; VALKOV; HINRICHS, 2014) that uses shadow projections on the display surface to create an interaction volume by the user's contact point on the virtual object.

Another approach for indirect techniques is to decouple the interaction surface from the visualization screen. Knoedel and Hachet (2011) adapted the RST technique using an external tablet for constrained manipulations of 3D objects on a computer screen. Ohnishi et al. (2012) used an external tablet to map its 2D input to the surface of an arbitrary 3D object. Simeone and Gellerseny (2015) compared the direct touch to the indirect touch in to control an object within a densely cluttered 3D environment while avoiding distractor objects, and in another work, when Simeone (2016) investigated the potential of indirect multitouch interaction as an alternative to direct touch for 3D stereoscopic displays.

Figure 2.2: (a) Indirect technique using the same surface for interaction and visualization, (b) with external surface.



(a) Void Shadows (b) Virtual Interaction Surface

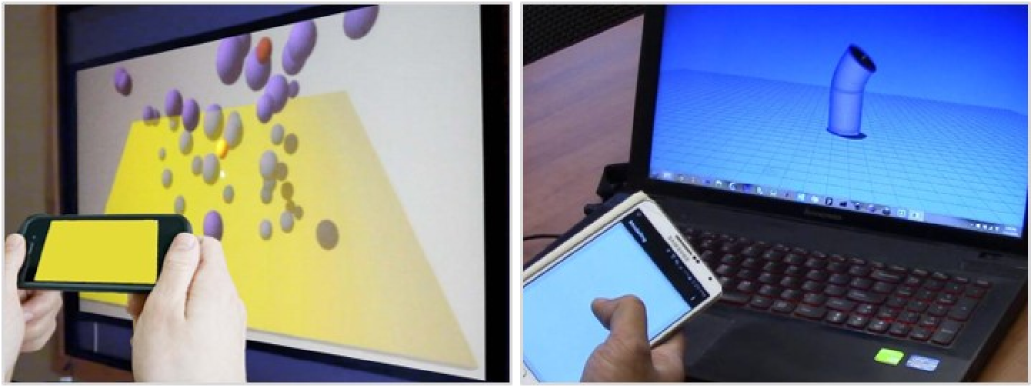
Source: (a) Giesler, Valkov and Hinrichs (2014), (b) Knoedel and Hachet (2011)

2.1.1.2 Inertial Sensors

Several interaction interfaces have inertial sensors embedded. They are essential to retrieve the device's orientation and acceleration. Commercial controllers such as the Wiimote revolutionized the way people interacted with video-games. The Wiimote introduced inertial sensors that allowed body actions to be mapped to similar actions in the game, such as, arm movements to interact with games like bowling and tennis.

Mobile devices are packed with inertial sensors that can be integrated into manipulation techniques for fast and precise orientation of 3D objects. Early work has explored the orientation of a mobile device to rotate an external object intuitively. Katzakis and Hori (2009) showed a gain in time with the 1:1 mapping of a device and virtual object orientation as compared to mouse and touchscreen input. Song et al. (2011) used the orientation of a mobile device to rotate a slicing plane in a medical volume data. The mobile device is placed close to a large projection, and the slicing plane extends from the edges of the device. The smartphone rotation is also explored in Tiltcasting (PIETROSZEK; WALLACE; LANK, 2015), which uses a virtual plane to select 3D objects in crowded virtual environments, and in MobiSweep (VINAYAK et al., 2016) for creating, modifying, and manipulating 3D shapes on a second screen. Katzakis et al. (2015) developed INSPECT a set of indirect touch techniques for 6-DOF manipulations using a smartphone. They use a plane-casting technique (KATZAKIS, 2012) for 3D positioning. Touch and slide gestures perform rotations. Their technique covers rotations and selection as well. They also performed an evaluation compared with an implementation of WAND, classical 6-DOF user interface. They report a significant advantage in translation time using the plane-casting

Figure 2.3: Techniques that use Inertial Sensors for manipulations. (a) Uses the inertial sensors to rotate a plane to facilitate selection in crowded environments, (b) uses rotations to create 3D shapes.



(a) Tiltcasting (b) Mobisweep

Source: (a) Pietroszek, Wallace and Lank (2015), (b) Vinayak et al. (2016)

when compared with a wand technique. Debarba et al. have explored inertial sensors combined with touchscreens to provide coarse + fine-grained 3D selection (DEBARBA; NEDEL; MACIEL, 2012; DEBARBA et al., 2013). The former reference also includes 2D translation and uniform scaling of distant objects.

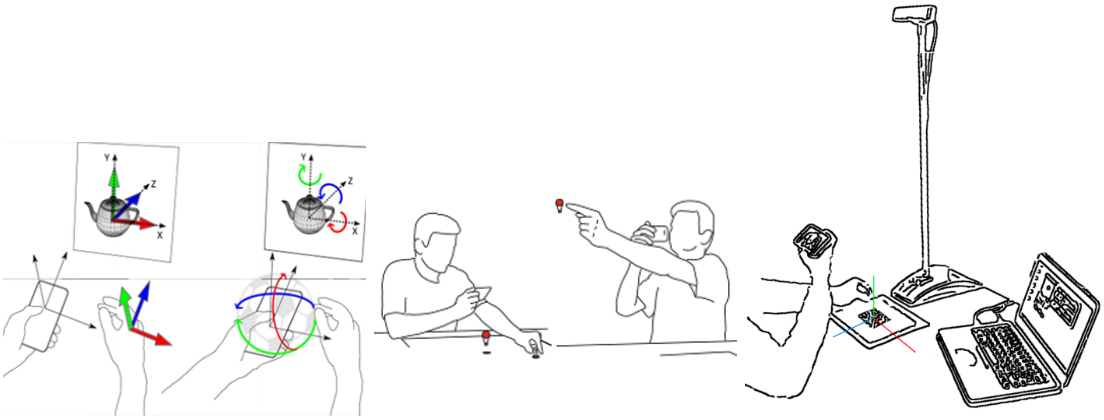
Liang et al. (2013) investigated the most natural sensory inputs to use in a tablet, and the most natural interactions to perform in such devices. Although the authors have evaluated the interactions with a tablet, they point out that smaller devices would perform better in motion gestures. A tablet was also the interface for Lopez et al. (2016) for their touch-based navigation of 3D visualizations in large stereoscopic displays.

2.1.1.3 Interaction Around the Device

Interaction around the device consists in perform gestures in the proximity of the device. Space is extended beyond the physical boundary of mobile devices to include the full 3D space around them (KRATZ; ROHS, 2009). The first *around the device* techniques used simple infrared (IR) sensors for multi-touch input in the area on the sides of the device for 2D object interactions (BUTLER; IZADI; HODGES, 2008) and hand gesture recognition above the device's screen (KRATZ; ROHS, 2009).

Precise manipulation of 3D elements requires better tracking apparatus than only the mobile device camera. It is usually achieved by attaching motion tracking cameras to the device or using external tracking systems. In *PalmSpace*, the authors implement mid-air gestures for 3D rotations of objects displayed on the device screen. The technique supports hand and fingers gestures without additional fiducials. This extra hardware allows

Figure 2.4: Around the device techniques. (a) Uses tracking with IR Cameras, (b) uses finger tracking and (c) uses depth camera.



(a) Around the Smartphone (b) Finger Tracking (c) Depth Camera

Source: (a) Bergé, Dubois and Raynal (2015), (b) Hürst and Van Wezel (2013) and (c) Bai et al. (2014)

interaction an input space larger than the arm's reach (KRATZ et al., 2012).

Bergé et al. (2014) have designed a smartphone-based overview+detail interface to interact with 3D public displays. They focus on solutions for the translation of the detailed smartphone view on the overview public screen. They evaluated three interaction techniques to control the movements of the detailed view: classical touchscreen pad, mid-air movements of the mobile device and around the device mid-air hand. In their experiments, mid-air techniques performed better than touchscreen input, and mid-air phone was the most suitable in general public conditions. In another work Bergé, Dubois and Raynal (2015) extend the mid-air hand into an "Around the smartphone" (ASP) approach to manipulate 3D elements on a second screen (Figure 2.4a). They evaluated the technique against an existing tangible and a tactile implementation in two user studies. Both ASP and tangible performed similar and better than the tactile technique. Rotation tasks were statistically better performed with the tangible technique. In this technique, the rotations follow the wrist movements of the user in a 1:1 mapping. They used a tracking device for rotation calculations instead of the smartphone sensors.

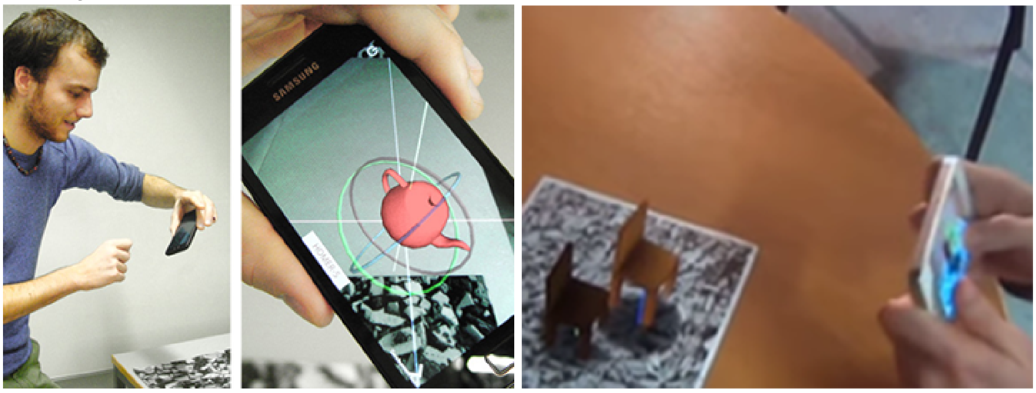
Gesture-based interactions around the mobile device are explored in augmented reality scenarios for objects interaction in the physical environments. Hürst and Van Wezel (2013) (Figure 2.4b) were some of the first to explore finger tracking for gesture-based transformations. Their setup only uses the default RGB camera and the inertial sensors of a smartphone for finger tracking and object positioning. They evaluated the user's performance and engagement for the canonical interactions, translation, rotation, and scale.

Even though the results turn the technique unfeasible for tasks that need precision, the concept achieved a high level of engagement and good performance when adapted for board games. It occurs mostly due to the lack of a robust finger tracking, such as the used by Bai et al. (2013), that attached a depth camera to a tablet. In another work, Bai et al. (2014) proposed a different setup where the depth camera is connected to a server and share the same coordinate system of the conventional mobile device camera through an AR marker (see Figure 2.4c). In this configuration, the server processes the finger tracking and send the information to the mobile device, thus, demanding less processing on the mobile side.

2.1.1.4 Device Movements

The advantage of manipulating 3D objects in augmented reality scenarios using mobile devices is the ability to use the physical movements for interaction. This approach is a natural way to place objects in the scene, as it mimics the real actions. It relies on the quality of the tracking for the amplitude of the movements, but on the other hand, it does not require an external apparatus attached to the device. Henrysson, Billinghurst and Ollila (2005) was the first to propose the use of movements to manipulate virtual objects. It was possible to alter the object's transformations by changing the device's pose. Samini and Palmerius (2016) designed a device movement technique that uses the user perspective rendering approach. The technique was compared with a fixed and relative device perspective approaches for near and far objects. Hybrid techniques take the advantages of the device movement and touch gestures. (MOSSEL; VENDITTI; KAUFMANN, 2013) proposed two techniques, *3DTouch* and *HOMER-S* for one-handed interactions with handheld devices (Figure 2.5a). The first implements touch gestures along with interface widgets to choose one transformation at a time. The *HOMER-S* works with the user's movements. The user first selects the action and then moves the device in the physical environment to transform the 3D object. During a manipulation task, the two techniques are integrated and can be combined. Similarly, Marzo et al. (MARZO; BOSSAVIT; HACHET, 2014) combine multi-touch and device movements. However, their interface is designed for two hands interaction (Figure 2.5b).

Figure 2.5: Techniques that explore the device movements as input for 3D transformations (a) is designed for one hand, (b) for two hands.



(a) 3DTouch + HOMER-S (b) Multi-touch + Device Movements

Source: (a) Mossel, Venditti and Kaufmann (2013), (b) Marzo, Bossavit and Hachet (2014)

2.1.2 6DOFs Tracked Controllers

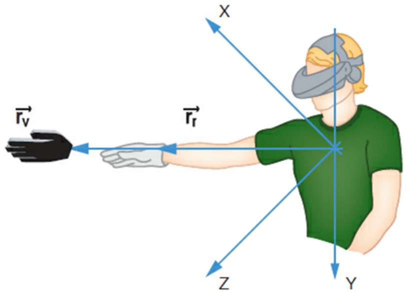
Manipulation techniques for 6DOFs tracked controllers became the standard interface in immersive VR with the popularity of Head-Mounted Displays. They are capable of mapping its position and orientation into the virtual space directly. Thus, a direct mapping of the user's hands to their virtual counterpart (*isomorphic*) is the conventional approach to interact with 3D objects. The direct mapping has significant shortcomings, such as tracking range restrictions and reaching distance. In the following sections, we compile *nonisomorphic* techniques design for 6DOFs controllers that minimize the disadvantages of the *isomorphic* approach.

2.1.2.1 Interaction at Distance

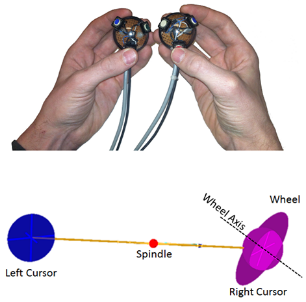
The interaction with objects beyond the arm's reach overcomes the restrictions of the interaction space present in the direct mapping. Poupyrev et al. (1998) proposed the *GO-GO* technique for distant interactions (Figure 2.6a). It uses a non-linear function to map the movement from the real hand to its virtual counterpart. For instance, when the user's hand is close to the user's body, *GO-GO* acts similarly to the direct approach. However, the mapping becomes nonlinear when the hand goes beyond a predefined distance, extending the virtual arm and allowing to access and manipulate distant objects.

In *Ray-casting* techniques the user points to an object to select it. Then, the object translates and rotates with the device while keeping an invariant rigid transformation relative to it. This technique suffers from several problems. First, it becomes challenging

Figure 2.6: Techniques to interact with distant objects (a) Go-GO , (b) Spindle + Wheel.



(a)



(b)

Source: (a) Poupyrev et al. (1998), (b) Cho and Wartell (2015)

to select objects that are far away. Some works solve this problem by using a cone-casting volume (FORSBERG; HERNDON; ZELEZNIK, 1996) or by disambiguation with progressive refinement (DEBARBA et al., 2013). Second, by maintaining a rigid transformation between the hand and the object, translations and rotations are coupled. Finally, the accuracy of movements varies according to the handling distance. The *HOMER* (hand-centered object manipulation extending ray-casting) technique proposed by (BOWMAN; HODGES, 1997) combine ray-casting with direct rotations to solve the inherent problems of *Ray-casting* techniques. For that, the user first selects the object using ray-casting. Then, the rotation changes to direct mapping.

2.1.2.2 Bimanual Interactions

Most techniques for 3D manipulations are designed for one hand interaction to avoid hardware availability limitations. The popularization of HMD headsets made bimanual designs practical by delivering two controllers, one for each hand. According to to Ulinski et al. (2009), bimanual interactions can be classified in symmetric, where both hands perform identical movements, and asymmetric, where each hand performs different movements.

Mapes and Moshell (1995) proposed an *Spindle* technique for symmetric bimanual interactions. The center of the two controllers represents the point of interaction that is used for both selection and manipulation. To translate an object both hands must move in the same direction. Rotations work by rotating the hands relative to each other. The *Spindle* technique can only perform yaw and roll rotations. The object's scale is modified by moving the hands closer or away from each other. Cho and Wartell (2015) addressed the *pitch* rotations, not present in the original *Spindle* using direct mapping of the dominant hand's rotations, Figure 2.6b. The technique is characterized as asymmetric due to the different actions both hands perform in rotations. Schultheis et al. (2012) reported that the *Spindle* is more efficient than the virtual hand when users are trained.

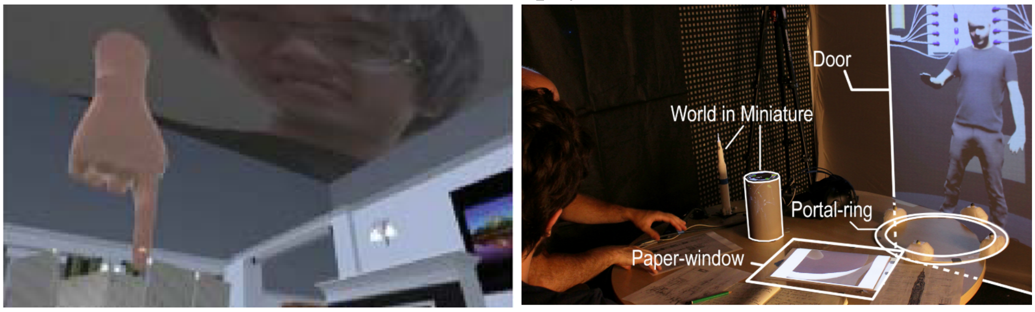
2.2 Asymmetric Interaction in Collaborative Mixed Reality Environments

In the context of collaborative virtual environments, asymmetry represents the capacity of individuals in a group to have different means to visualize and interact with virtual content. The technique created by Billinghurst et al. (BILLINGHURST et al., 2000) is the early system that mixes different interaction and visualization concepts for co-located VR and AR asymmentric collaboration. Generally, asymmetry can vary in scale, roles, realities and devices, and depending on the design they are linked together.

Asymmetry in scale allows users to have different sizes in the VE, and consequently, allowing for different Points-of-View (PoV), see Figure 2.7. Besides the human scale that is the standard, bigger scales may be represented as an analogy for a giant or a god and smaller scales for an ant. Due to the particularities of each perspective, the participants may have different roles and use different devices, controls and displays to execute the tasks. The concept of different scales in VEs was first introduced with the world-in-miniature (WIM) technique (STOAKLEY; CONWAY; PAUSCH, 1995). In the original version, a single user could alternate between and scales to interact with virtual objects. In multiuser collaborative VEs asymmetric scales are assigned to different users and explored in navigation (NGUYEN; DUVAL; FLEURY, 2013; WANG et al., 2012), manipulation of 3D objects (CHENECHAL et al., 2016), data exploration (SUNDÉN; LUNDGREN; YNNERMAN, 2017) and architectural design (GOLDSCHWENDT et al., 2014; IBAYASHI et al., 2015). The work proposed by Clergeaud et al. (2017) mix virtual and augmented realities to improve interaction, awareness and communication between remotely located participants, see Figure 2.7b.

CVEs can be designed to handle asymmetric interactions while users keep the same scale. Those systems usually support interaction with different input devices and visualization hardware while preserving its usability across the asymmetric modalities. ShareVR (GUGENHEIMER et al., 2017) combines *HMD* users with *Non-HMD* users into a spatial augmented reality and virtual reality environment. The prototype enables co-located users to interact with each other in the same shared environment. The authors report an increase in enjoyment, presence and social interaction when compared with the

Figure 2.7: Asymmetric interaction in collaborative mixed reality environments. (a) Dollhouse , (b) Seamless interaction between physical and virtual locations.



(a) (b)

Source: (a) Ibayashi et al. (2015), (b) Clergeaud et al. (2017)

Non-HMD user interacting in a conventional TV+Gamepad setup. In the same direction, CoVAR (PIUMSOMBOON et al., 2017) implements interactions between AR and VR. The AR user scans and shares its local environment with the VR user while they are in remote locations to increase awareness of the space during the interactions.

The popularity of HMDs brought the asymmetric metaphor to commercial games, such as, in the *Playroom VR, Black Hat Cooperative and Ruckus Ridge VR Party*, where a user wears the HMD and has a different perspective from players who are using the conventional game controllers and screen.

Similarly to prior works, we allow users to cooperate using asymmetric interactions. However, in our solution users can share the transformation tasks in simultaneous fluid manipulation of a same 3D object. Users decide whether to block or divide the actions. Thus, groups can adopt their strategy without system limitations. In Chapter 6, we investigate the effect of simultaneous manipulations in an asymmetric setup.

2.3 Collaboration in Virtual Environments

Computer-Supported Cooperative Work (CSCW) is the field that focuses on the collaborative aspects of human interactions through computers. Collaborative systems allow taking advantage of the skills from each collaborator to solve complex tasks. It is a cross-field that can benefit any digital area that the work is centered around the human. The CSCW has coined the main aspects of human collaborative work. While CSCW is broad enough to coined the main aspects of human collaborative work, it is mostly centered in interactions in traditional desktop computers (groupware). Collaborative Virtual Environments (CVEs) are shared spaces designed to support interactions between users and objects in a virtual world (VR) (CHURCHILL; SNOWDON, 1998). Different from groupware applications, in CVEs, users have more degrees of control and more types of interactions with digital objects. We compiled a list of aspects that we consider important for collaborative VEs. The list was based on the classifications present in (MARGERY; ARNALDI; PLOUZEAU, 1999; OTTO; ROBERTS; WOLFF, 2006; AGUERRECHE; DUVAL; LÉCUYER, 2009; DODDS, 2009; NGUYEN; DUVAL, 2014). We emphasize the collaborative features that are essential in immersive virtual environments.

Awareness: Awareness is the understanding of both personal and shared task-related activities (DOURISH; BELLOTTI, 1992). This understanding allows individuals to evaluate their actions, group goals and work progress. Awareness is essential

Co-Presence: The users recognize the co-presence through a representation of other members in the virtual environment. Avatars are the most common representation in VR, but it can vary depending on the system. The co-presence is perceived when the modifications in attributes of objects or the scene are visible to all collaborators (NGUYEN; DUVAL, 2014).

Active/Passive: The users in a system can be active or passive. Active users are those who control the interaction and the manipulation process. Passive users cannot perform actions. This approach is common in CAVEs where all users are immersed with stereoscopic glasses, but only one has the control over the interaction.

Synchronous/Asynchronous: In synchronous CVEs, two or more users can interact, handle and manipulate the same object. While in asynchronous, only one user can change the proprieties of an object at a time.

Symmetrical/Asymmetrical: Symmetrical collaborators benefit from similar viewpoints and interaction capabilities. While in asymmetrical the users collaborate with different points of view, roles or devices.

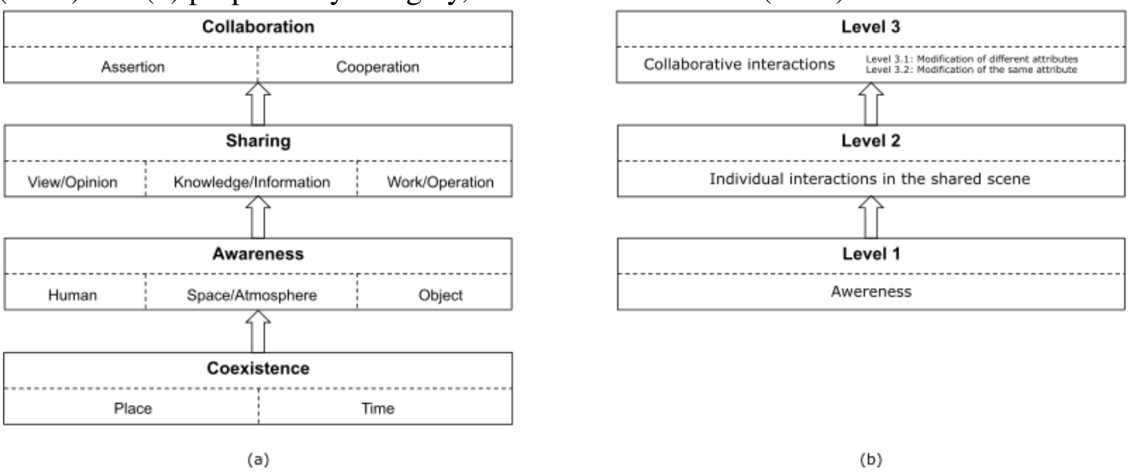
Co-located/Remote: Wheather the users are in the same physical environment or are distributed in different locations while sharing the same virtual environment.

Other important aspects in collaborative virtual environments are communication and network latency. While they are crucial for reliable interactions in VEs, it is not the focus of this work. An in-depth survey about network latency in VEs is presented by Khalid, Ullah and Alam (2016).

2.3.1 Collaborative Interactions in VEs

The demand for spatial collaborative virtual manipulations emerges when single-user tasks become too difficult to perform due to the many degrees of freedom involved and when the task requires the participation of multiple users. It may happen in many application fields, such as simulation and training, data exploration (RUDDLE; SAVAGE; JONES, 2002), industrial manipulations (AGUERRECHE; DUVAL; LÉCUYER, 2010; SALZMANN, H.; ACOBS; FROEHLICH, 2009), and exploration of large environments (NGUYEN; DUVAL; FLEURY, 2013). The works of Margery, Arnaldi and Plouzeau (1999), Okada (2007) classify the collaborative interactions based on the features provided to the user for cooperation. These classifications are hierarchical, ranging from basic to advanced features. Figure 2.8 shows the hierarchical organization created both authors. We find the Margery, Arnaldi and Plouzeau (1999) classification more expressive for the context of this thesis since the authors expand the collaborative manipulations in the modification of different attributes (Level 3.1) and modification of the same attribute (Level 3.2).

Figure 2.8: Hierarchical classification of collaborative interactions. (a) proposed by Okada (2007) and (b) proposed by Margery, Arnaldi and Plouzeau (1999).



Source: the Author, based on (OKADA, 2007)

2.3.2 Co-manipulation in Collaborative VEs

Until Level 2 of collaborative interactions (MARGERY; ARNALDI; PLOUZEAU, 1999), elements on the scene are controlled by only one entity at a time by each user. In such way, the transformation desired is directly calculated and applied in the object by matrix multiplication in the form of $M = MS$, where M represents the object's matrix and the S is the new transform, both have $4x4$ dimension. The S matrix represents the three fundamental transformations: translation, rotation and scale.

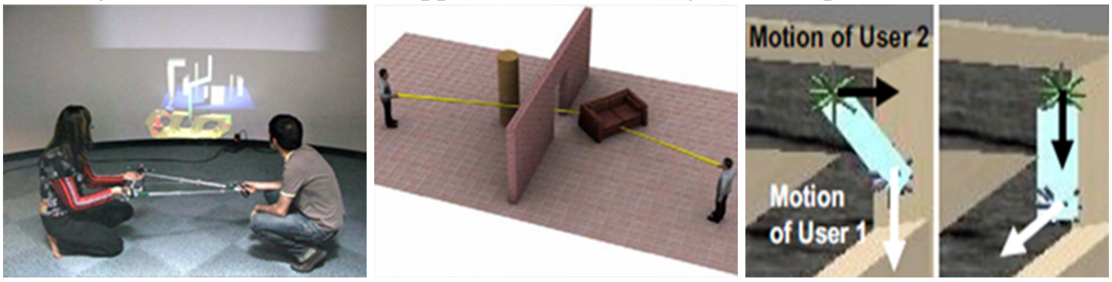
Problems start to appear when two different sources try to change the same object simultaneously (Level 3). In this case, both sources S_1 and S_2 would change their version of the object M_1 and M_2 respectively and then update the shared object, $M = M_1S_1$ and $M = M_2S_2$. As a result, the final object's M transform would be that applied for the last in time. For instance, if S_1 set a 45° rotation in time $t1$ and S_2 set a -90° in time $t2$ the final object's rotation would be -90°. Moreover, the object would leap between the changes causing an undesired visual flickering. In the end, the contribution made by $S1$ is overwritten by $S2$ as it never occurred.

The desired behavior would represent the contributions of each source in some way. Thus, one of the main goals of this thesis is the design of simultaneous collaborative interactions in virtual environments. We focus on works that explore the synchronous approach that corresponds to the *Level 3.2* of cooperation proposed by Margery, Arnaldi and Plouzeau (1999). Following, we present the main approaches that were proposed in previous works

Different from the real world, where a physical link is created between the users that are carrying an object, in the virtual world this link does not exists. The lack of a physical link between the users and the object in VEs is the main limitation present in virtual co-manipulation and can affect the effectiveness of manipulations. Thus, techniques that use a tangible user interface (TUI) provide a physical form to control virtual objects. A TUI can be non-reconfigurable – the physical object cannot be modified, or configurable, where the shape can change. Aguerreche, Duval and Lécuyer (2010) designed a configurable tangible device (CTD) that can be manipulated by single or multiple users to control a virtual object in a VE (Fig. 2.9a). Similarly, SALZMANN, H., ACOBS and FROEHLICH (2009) propose to use a non-reconfigurable tangible device for two-user collaborative interactions. They also compare their technique with a non-tangible method in an assembly task. In both works, the authors report that a prop-based interaction improves task performance and collaboration because of the link between the two users provided by the tangible device. TUIs are fundamentally designed for co-located due to the need of a physical prop. Another problem is that TUI is rather limited by the shape of the prop that can only represent a set of virtual objects even in the configurable approach.

Several approaches were proposed to combine users' movements for virtual objects

Figure 2.9: Collaborative Approaches for 3D Objects Manipulations in VEs.



(a) Tangible Device (b) DOF Separation (c) SkeweR Technique

Source: (a) Aguerreche, Duval and Lécuyer (2010), (b) Pinho, Bowman and Freitas (2008) and (c) Duval, Lécuyer and Thomas (2006)

positioning without the need of a TUI. Ruddle, Savage and Jones (2002) classify the techniques into the symmetric integration of movements, where users perform coordinated actions (e.g. both users move the object up), and asymmetric integration of movements, where users can perform different actions on the object (e.g. one user rotates another user moves). The authors compare the symmetric and asymmetric approaches in a study where users need to carry a large object through a corridor. They report that the symmetric approach is more relevant when the two users have to perform a very similar action, while the asymmetric approach is preferred when users have to perform different tasks. In VR, where it is possible to extrapolate the bounds of the real to enhance user's abilities, an asymmetrical approach is more practical for co-manipulations in VR since the users' movements are not dependent on each other's. On the other hand, since the actions are only possible in virtual environments, users may need more time to be comfortable with the interaction.

The Bent-Pick-Ray (RIEGE et al., 2006) is an asymmetric technique for co-manipulation in VR based on the ray-casting metaphor. When two users are manipulating the same object, the selection rays bent based on the pointing direction and the point of selection. This visual feedback helps users understand each other's actions since manipulations are simultaneous. Rotations are computed with a spherical linear interpolation and translations are interpolated using only the difference of transformations. DOF separation consists in splitting the tasks among users. In this case, the number of DOFs that each user can access and control is limited: one user controls the rotation of the object, while the other one is limited to translation. Pinho, Bowman and Freitas (2008) explore this approach to demonstrate that the use of cooperative interaction techniques can be more efficient than two users working in parallel using single-user interaction (Fig. 2.9b). The SkeweR (DUVAL; LÉCUYER; THOMAS, 2006) technique enables multiple users to simultaneously grab any part of a virtual object through points called "crushing points". To determine the translation and the rotation of a grabbed object, SkeweR considers positions of these points and average the 3D object final position (Fig. 2.9c).

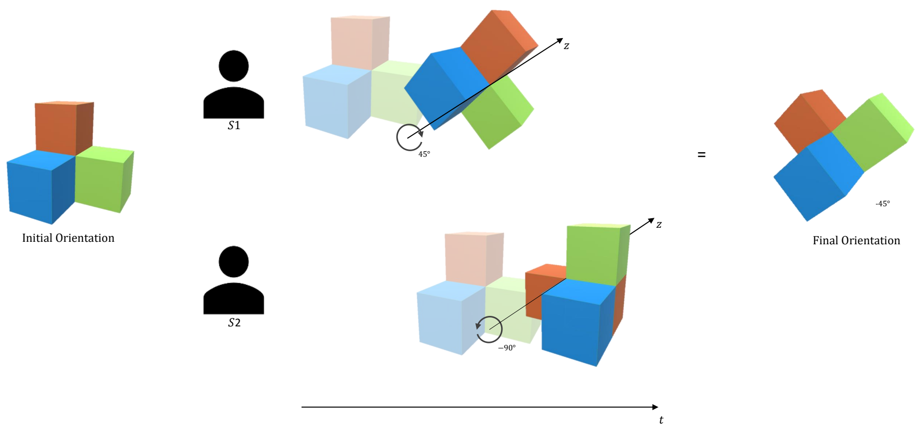
2.4 Literature Discussion

We organized this chapter putting in evidence the topics more closely related to the remainder of this thesis. The understanding of the strengths and weaknesses of the previous methods guides our design decisions and allows us to seek for unsolved problems in the field. One contribution envisioned in this thesis is to explore collaborative aspects for *simultaneous* object manipulation. Hence, we have researched works that apply group work to solve manipulation tasks. We have found that co-manipulation techniques are designed for specific applications, hardware or visualization. Furthermore, all previous works use other interfaces than mobile devices and do not scale well for multiples users in simultaneous object manipulations. Most works are evaluated for two-user scenarios only, leaving open questions about the collaboration between three or more users. We took these shortcomings into account during the development of this thesis. Now, we briefly introduce the concept of our solution for the co-manipulation problem. In the following chapters, we detail all design choices and evaluation methods we carried out towards seamless and efficient simultaneous collaborative 3D interactions and their application on virtual, augmented and mixed reality.

2.4.1 Introducing the *Sum-of-Contributions* Approach

In pursuit of a more general form to represent concurrent access to the same DOF, we designed the *Sum-of-Contributions* (SoC) model. The purpose of SoC is to provide a simple but robust approach for simultaneous manipulation of virtual objects. SoC is designed to seamlessly aggregate inputs from multiple sources directly and unrestrictedly. Thus, every action performed by each source counts in the final object's transformation. For instance, in the previous example where a 45° and a -90° rotations are applied by S_1 and S_2 respectively, the final object's rotation would be -45°, as shown in the Figure 2.10. Therefore, if $S1$ and $S2$ rotate the object in exactly opposite directions, the action are canceled and the rotation of the object will not change. If the sources manipulate different

Figure 2.10: Simultaneous access to the same DOF when two sources are trying to change the object's orientation. The *Sum-of-Contributions* model calculates the final object orientation.



Source: the Author

transformations in parallel, the inputs are merged.

The SoC uses transform increments to change the object's position, rotation and scale. Those increments are the difference between an input S in frame f and an input in frame $f-1$. We call this difference a *Transformation Step* and it is described as S_{step} = SS^{-1}. Finally, the object receives the *Transformation Step* contribution: M = MS_{step}. The extension for multiple input sources is straightforward. It is only necessary to include the contributions of the different sources: $M = MS1_{step}S2_{step}...Sn_{step}$

The *Sum-of-Contributions* characteristics are:

- As SoC is a generic model for simultaneous manipulations, it does not implement any restrictions or priorities on how the transformations are applied. The responsibility of these elements is on the interface design.

- The inputs from several sources are integrated asymmetrically. Thus, users can execute different action on the same object.

- The asymmetry leads to more complex strategies where users can choose to split the manipulation tasks. On the contrary of DOF separation approaches, where the separation is enforced.

- The transition from single-user to collaborative is seamless, without any mode switch or restrictions on users.

The SoC formulation presented is a robust yet straightforward model for simultaneous manipulations. As discussed above, the model has to be integrated into the 3D user interface design to reveal its benefits. In the Chapters 3, 4, 5 and 6, we apply the SoC in the design of novel user interfaces for simultaneous collaborative interactions with 3D virtual objects. We prove the versatility of the model in an extensive evaluation in a variety of immersion scenarios, tasks, hardware and group arrangements.

3 HANDHELD-BASED 3D USER INTERFACE FOR COLLABORATIVE OBJECT MANIPULATION IN VIRTUAL ENVIRONMENTS

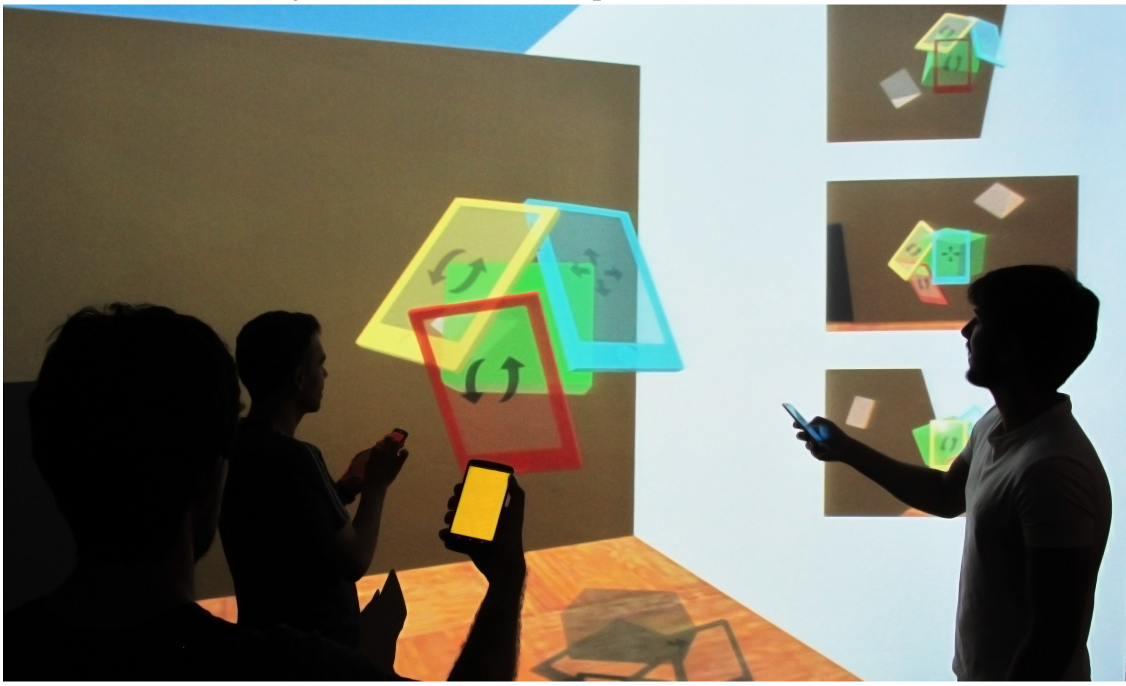
As discussed in the *Literature Review*, assessments that investigate collaboration with different group sizes are limited to the maximum of two members. Thus, we aim to investigate the relationship between group sizes and the time and accuracy to complete tasks in VE. Furthermore, we intend to understand the influence of work distribution balance and work division in the performance of group combinations. For that, we describe a novel 3DUI for collaborative object manipulation. Our technique explores the advantages of smartphones as an interface for 3D interaction. We use the inertial sensors, the touchscreen and the physical volume buttons to manipulate a total of 7 DOFs of a selected 3D object. More specifically, there are 3 DOFs for translation, 3 DOFs for rotation and 1 DOF for uniform scale. The transformation actions we exploit are consolidated in the everyday tasks when interacting with mobile apps and games, such as touch and slide to translate, device orientation to rotate, and pinch and spread to scale the virtual object. Our approach accepts simultaneous connection of multiple smartphones, allowing users to collaborate by simultaneously controlling any manipulation DOFs. This produces equal participation, individual responsibility and positive interdependence, which are necessary features in a collaborative task.

3.1 3D Manipulation Interface

3.1.1 Graphical Representation

The representation of each mobile phone in the VE includes two elements: (i) a smartphone 3D model that orbits around the selected object following the phone orientation (Figure 3.2). When the user rotates the device, the virtual cursor (virtual representation of the smartphone) follows, assuming the same orientation and (ii) a picture-in-picture (PIP) camera containing the point of view of the orbiting phone model, positioned at the top right corner of the display (Figure 3.1). The PIP camera is meant to provide additional depth cues, which are rather limited when the display is shared among multiple users. Each connected device is labeled with a color. The user interface includes a global camera and display, which are shared by all collaborating users. All mobile phones are calibrated with

Figure 3.1: Three users simultaneously driving the object through the virtual environment. The colored rectangles indicate the position and orientation of each user in the VE. The three windows at the right-side show the three personal views.



Source: the Author

respect to the screen position.

We also render icons on the virtual phone representation indicating the transformation being performed by each user, so that collaborating users can be aware of each other's actions without the need of verbal communication. The icons can indicate that a mobile phone is either performing a translation, rotation or scale on the selected object (Figure 3.2a-c).

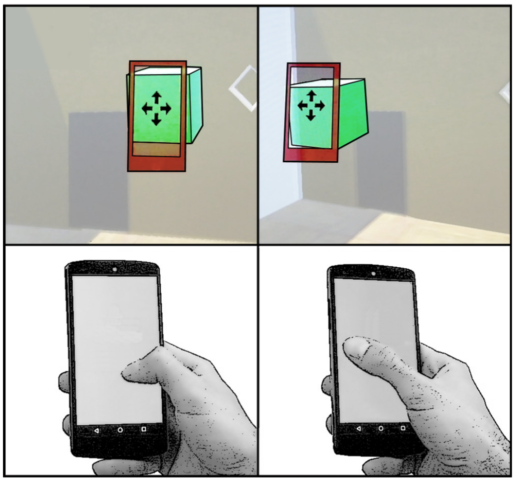
We chose to not use the device display for visualization. This avoids the focus changing between the main screen and the device display that can lead to the losing of context.

3.1.2 Manipulation Mapping

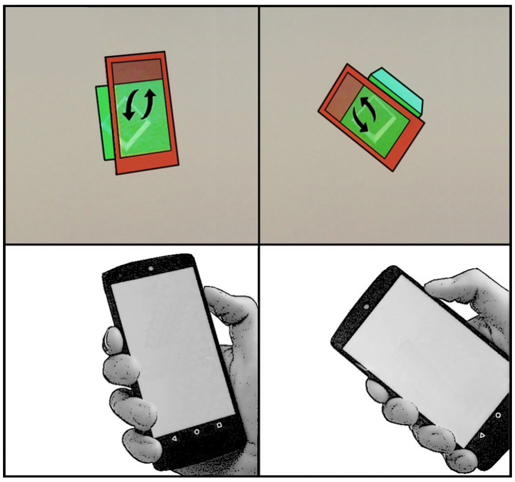
Our 3D manipulation interface uses the mobile phone physical orientation, the touchscreen, and the physical volume buttons to manipulate a total of 7 DOFs of a selected 3D object. More specifically, there are 3 DOFs for translation, 3 DOFs for rotation and 1 DOF for uniform scale.

We use the sensors fusion capability provided by the Android API to retrieve

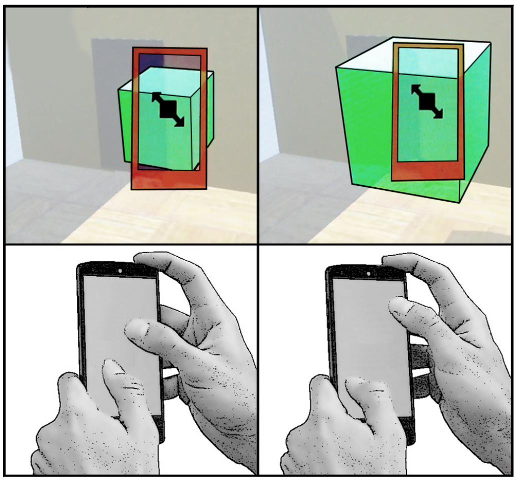
Figure 3.2: Walkthrough of manipulations that a mobile phone cursor can perform over the selected object: (a) translation can be applied by touching the mobile phone screen and sliding the point of contact, the object translates on the plane defined by the mobile device orientation; (b) by holding the volume down button and rotating the phone one can rotate the object likewise; (c) scale is applied by touching the screen with two fingers and producing a pinch/spread gesture; (d) the camera orientation can be controlled by holding the volume up button and rotating the mobile phone.



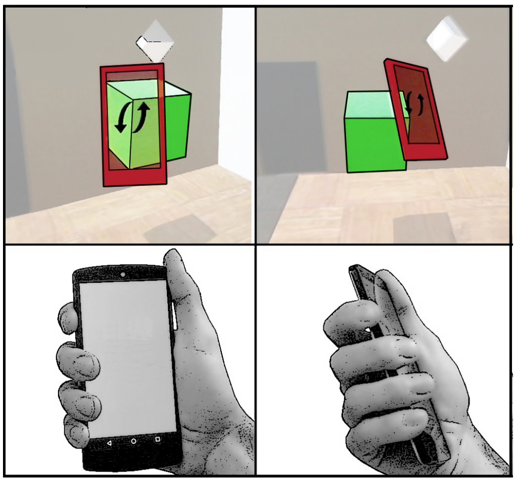
(a) Object translation



(b) Object rotation



(c) Object scale



(d) Camera rotation

Source: the Author

the device orientation. The fusion uses the accelerometer, gyro, and magnetometer to provide an orientation in the 3D space. A calibration action is used once to provide a 1:1 mapping. To translate the selected object, a plane is defined using the device orientation (1:1 mapping). This plane is aligned to the touchscreen and, as the thumb (or any other finger) starts to touch and slide on the screen, a corresponding translation is applied to the object (Figure 3.2a). The object translation rate is constant and equivalent to 1 unit in the VE for each 333 pixels of sliding over the device's screen. The cube has initial side of 1 unit. The selected object is rotated by pressing and holding the volume down button. Throughout the time the button is pressed, the virtual object will rotate following the orientation of the phone (Figure 3.2b). Thus, during a rotation action, the manipulated object follows a 1:1 mapping with the physical device. We do not allow changes in the rotation rate in order to preserve this absolute mapping. Clutch can be usedto reach a total rotation beyond wrist limits (i.e. perform object rotation, reposition the mobile phone, then perform a new object rotation). Uniform scale is performed by pinch and spread gestures with two fingers on the touchscreen (Figure 3.2c). The scale is obtained by dividing the current distance between the fingers by their initial distance. Finally, the main camera (shared among users) follows the selected object at a fixed distance. If a user wishes to change the camera orientation, this can be done using the volume up button in a similar manner as the virtual object rotation action (Figure 3.2d). The virtual shared camera can be only rotated in the pitch and yaw axes. If users wish a personalized point of view, they can use the embedded PIP camera.

A user can perform any transformation action isolated or two by two: rotation and translation, scale and rotation, camera rotation and translation, scale and camera rotation.

3.1.3 Generalizing for Collaborative 3D Manipulation

We do not impose limits on the number of simultaneous users. All users have access to all available functions. While working in groups, collaborating users may wish to split the different aspects of the manipulation among them. For instance, while the first user translates an object, the second user could rotate this same object.

We use the *Sum-of-Contributions* presented in section 2.4.1 to handle co-manipulation. Thus, every action performed by each individual user counts as a transformation step. Action matrices are multiplied by the transformation matrix of the virtual object. Therefore, every contribution from each user is summed up into the final object's transformation without restrictions or weights. Therefore, if two users move the object in opposite directions, the position of the object will not change (Figure 3.3a). On the other hand, if they manipulate the different transformations in parallel, the transformations are merged (Figure 3.3b). Figure 3.3 shows the concurrent access of same transformations by three users.

Since all the users' actions are applied directly to the manipulated virtual object, we smooth the movements with an exponential moving average filter (SMITH, 2013) to minimize the undesired flickering. This filter has been chosen because it only needs the previous values to filter the signal. In this work, the filter is in the format:

$$NewValue = PrevValue * (1 - C) + CurrentSignal * C, \quad (3.1)$$

Where C is the filter constant. We fixed $C = 0.5$.

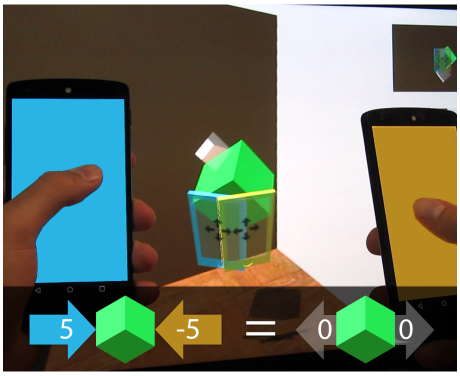
Regarding visualization, we chose to use a large screen as a shared visualization platform. It is a handy solution to aggregate multiple users in a single VE. Any person passing by the display is able to connect to the application and interact with it. This setup corroborates with the objectives of interactive public displays. Depending on the application goals, our technique can be easily adapted to HMDs, CAVEs and other immersive displays. Individual displays allowing parallax and binocular stereo have the potential to increase depth perception and user performance.

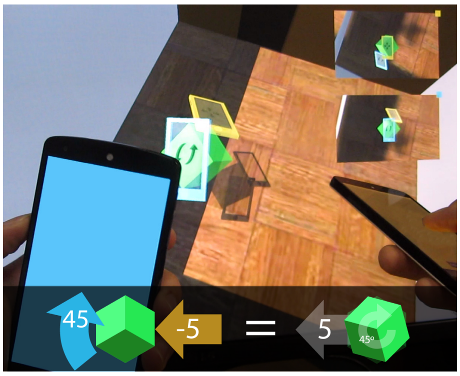
3.2 Collaborative 3D Manipulation Assessment

We conducted four public demonstrations of our collaborative user interface prior to the user study presented below. In these demonstrations we proposed different 3D manipulation tasks: object docking, and obstacle crossing. In the former, the users had to manipulate 7-DOF of selected objects in order to stack them by precisely docking one on top of the other. In the latter, the users had to manipulate a selected cube to take it from an initial to a destination position. The cube had to be carried through a sequence of walls with openings while avoiding collisions but occupying the maximum volume possible. Surface constraints were used to prevent the interpenetration between the cube and the obstacles. Over a hundred people have tested our technique in the demonstrations. We collected users' feedback to refine the interaction, ergonomics and user experience.

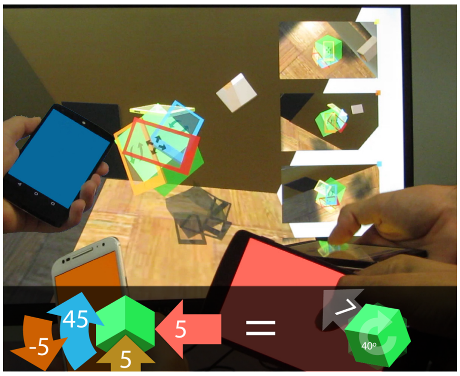
The obstacle crossing task was then used for a formal experiment to assess the

Figure 3.3: Concurrent access to transformations. Every action performed by each user counts as a transformation step. At the bottom of the image we show the applied forces and the resulting transformation.





(a) Forces cancellation (b) Transformations merge



(c) Concurrent access to the same DOF

Source: the Author

effect of group size on manipulation time and accuracy. Three tasks are part of the obstacle crossing experiment. The first task is a wall with a square opening scaled down but aligned to the axes of the manipulated cube. The second is a wall with a both scaled and rotated square opening. The third is a winding tunnel with cross sections that vary in width and orientation (Figure 3.4). Each task starts after a three seconds count down and finishes when the object reaches the destination area at the end of the obstacle. The first two walls are used for training and the tunnel is used for the evaluation.

The evaluation consists in calculating the time and the accuracy to transpose the tunnel. We calculated the time from the start to the end of the task and the accuracy through a checkpoint system. The checkpoints are objects identical to the manipulated object (yellow cubes in Figure 3.4). We placed eight checkpoints along the tunnel path with a static position, orientation and scale so as to occupy the maximum volume at that location. To perform the task with high accuracy the users had to occupy the maximum volume of these checkpoints with the manipulated cube while completing the circuit. To calculate the error between the manipulated object m and a checkpoint c, we take the minimum distance from a vertex m_i to all vertexes in c. Thus, the error for each checkpoint is the maximum distance among the minimum distances previously calculated. For each task, we compute the error median for the eight checkpoints.

To the best of our knowledge, there is no well established interface for 3D collaborative manipulation (translation, rotation and scaling) to permit a comparison with our technique. We are aware of slightly distinct input-output mappings of mobile input for some of these transformations in a single user mode, though, from very good results regarding precision and comfort (DEBARBA; NEDEL; MACIEL, 2012; DEBARBA et al., 2013). We used these previous results to justify our design choices and decided to focus on the evaluation of the collaborative aspects of our technique.

3.3 User Study

3.3.1 Design and Procedure

We aim to investigate the relationship between group sizes and the time and accuracy to complete the tasks. Furthermore, we intend to understand the influence of work distribution balance and work division in the performance of each group combination. Thus, the experiment follows a between-subjects design with *Group size* as the only independent variable, with one, two, three or four participants. Dependent variables collected were *time* to complete the task and *accuracy* of the group, and *transformation actions* (translation, rotation, scale or camera rotation), including duration and magnitude of the action performed by each individual subject. The accuracy is measured as described before in Section *Collaborative 3D Manipulation Assessment*.

Participants read and agreed to an Informed Consent Form (see Appx. C) and answered a characterization form before the experiment. They also watched an informative 2 minutes video about the technique and the task. Then, we handed the mobile devices already connected in the virtual environment to the participants. We demonstrated the calibration steps and, after that, we started with the training session. The training session began with an individual exercise and continued with the group practice. The session finished when all participants were satisfied, without time restrictions. Then, users recalibrated the devices (when needed) and we started the test. Between the trials, users could recalibrate the devices as well. We asked participants to prioritize accuracy over time. The test ends with a post-experiment questionnaire. In average, the experiment sessions lasted

30 minutes.

We first investigate the following hypotheses:

- H1. Groups with more than one member complete the tasks faster
- H2. Groups with more than one member complete the tasks with more accuracy
- H3. For the tested group size range, if groups increase in members, the time to complete tasks drops proportionally
- H4. For the tested group size range, if groups increase in members, the accuracy to complete tasks increase proportionally

If the hypotheses are confirmed, we can assume that our interface provides a significant gain in performance through collaborative work. Then, we can explore sub-hypotheses based on the first results.

3.3.2 Task

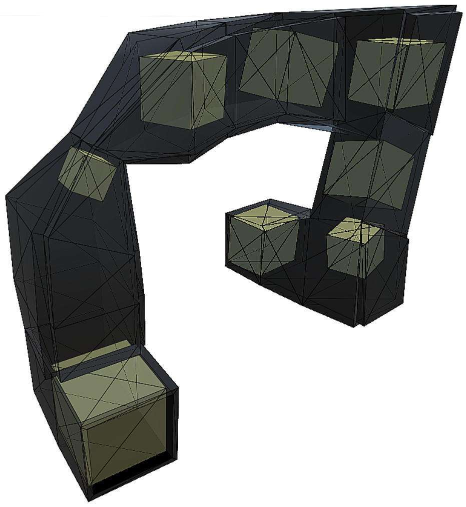
We used the obstacle crossing game with three wall configurations, as described in Section *Collaborative 3D Manipulation Assessment*. The training sessions consist of the first two walls. The test session is formed by one trial for each practice wall and two trials for the tunnel. In all trials, the selected object starts 4 units distant from the walls. After the obstacle, the object has to be carried to a finishing wall distant 4 units. We provide feedback information whent the object collided with walls.

The Section *Results* reports only the two trials in the tunnel task for the statistical analysis.

3.3.3 Subjects

Sixty subjects participated voluntarily in this experiment (nine female), aged 24 years in average (SD=3.6). They were all Computer Science students with no movement restrictions on wrists and arms. Thirteen of the individuals had never used gestural interactions with Kinect, Wiimote or mobile devices. We arranged the participants in 5 groups of one, 7 groups of two, 7 groups of three and 5 groups of four individuals.

Figure 3.4: Checkpoints in the obstacle crossing task.



Source: the Author

3.3.4 Experimental Setup

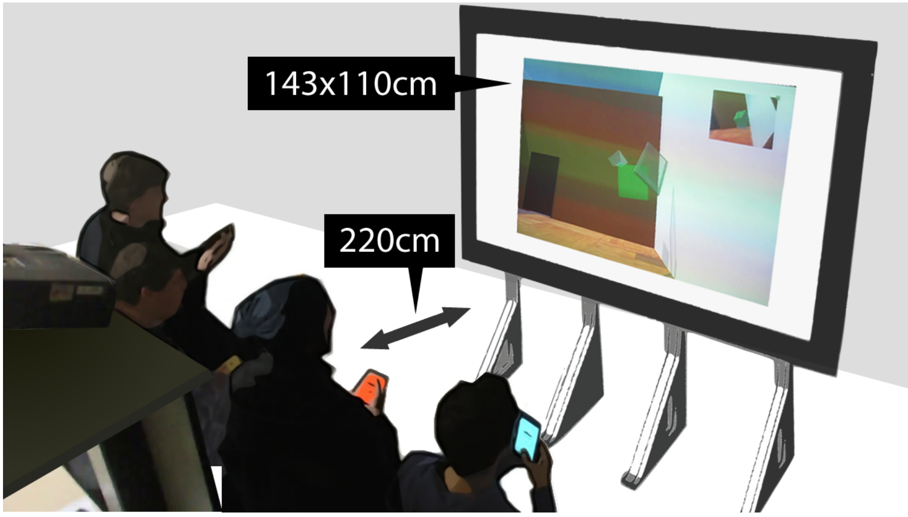
Our setup is composed of mobile devices, a server computer, a projection screen and a WiFi router. An app on the phone communicates with the server that manages the client's data and the virtual environment. In the experimental setup, the same screen was shared by all the users.

Although we used diverse handheld devices, all of them were Android-based smartphones with the 5.0+ version of the operating system and had WiFi connection. They had physical volume buttons placed on the side of the device, touch screen, and gyroscope and accelerometer sensors. The same compilation of the application was installed on all devices.

The server application runs on a PC and is developed on Unity 3D. It renders the virtual environment, manages the connections with the mobile phones, and the manipulation of the transformations to be performed. Communication between server and phones is via a dedicated WiFi router and uses TCP network protocol to ensure that the packets will arrive in the correct order.

The projector is a Sharp with 1024x768 resolution. It is placed at 225 cm above the participants and 260 cm far from the projection screen. The projected image has 143x110 cm in size. The participants are placed next to each other at a distance of 220 cm from the screen. Figure 3.5 shows the physical space setup.

Figure 3.5: Physical setup of the experiments. Participants are placed next to each other at a distance of 220 cm from the screen. Each of them has its own device.



Source: the Author

3.4 Results

3.4.1 Group Size vs. Time and Accuracy

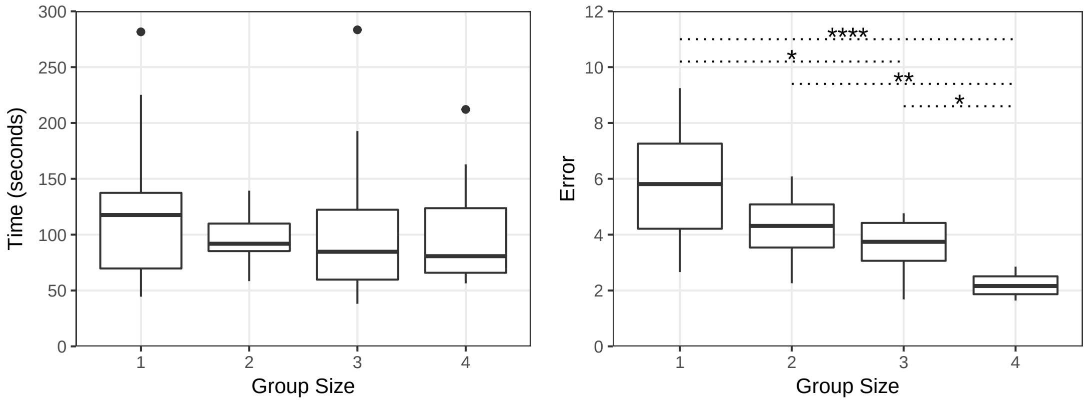
Figure 3.6 shows the time and accuracy that each group configuration achieved in the evaluation. For the statistical analysis, we first verified if the relation between the independent (group size) and the dependent variables (time and accuracy) could be evaluated using ANOVA. We fitted a linear regression and tested the normality of the residuals (Shapiro-Wilk test). Residuals were not normally distributed on both comparisons. Although Norman (NORMAN, 2010) suggests that ANOVA is robust when residuals are not normally distributed, we made a conservative choice by using non-parametric tests as these make claims about the difference of medians instead of the average.

We conducted a Kruskal-Wallis test, which can determine if a statistically significant difference exists between the levels of *group size* without assuming residuals to follow the normal distribution. Post-hoc was conducted using paired Dunn tests with Holm-Bonferroni correction.

3.4.1.1 Group Size vs. Task Completion Time

The Kruskal-Wallis test failed to reject equality of medians across different group sizes for task completion time ($H(3) = 2.1834$, $p = 0.54$), thus we reject H1 and H3. See Figure 3.6a.

Figure 3.6: (a) There was no significant difference in time to complete the tasks between groups. (b) Group size 4 is the most accurate (Mdn = 2.16, SD = 0.44), followed by size 3 (Mdn = 3.74, SD = 0.94), size 2 (Mdn = 4.31, SD = 1.18), and size 1 (Mdn = 5.81, SD = 2.16).



(a) (b)

Source: the Author

3.4.1.2 Group Size vs. Task Accuracy

As we hypothesized in H2, the Kruskal-Wallis test revealed significant effect of group size on task accuracy median ($H(3) = 21.3522$, $p < 0.0001$). The post-hoc Dunn test indicates that significant accuracy increase occurs between group sizes 1 and 3 ($p = 0.0486$), 1 and 4 ($p < 0.0001$), 2 and 4 ($p = 0.0024$) and 3 and 4 ($p = 0.0226$). This result confirms H4. No significant difference was measured between group sizes 1 and 2, and between group sizes 2 and 3. See Figure 3.6b.

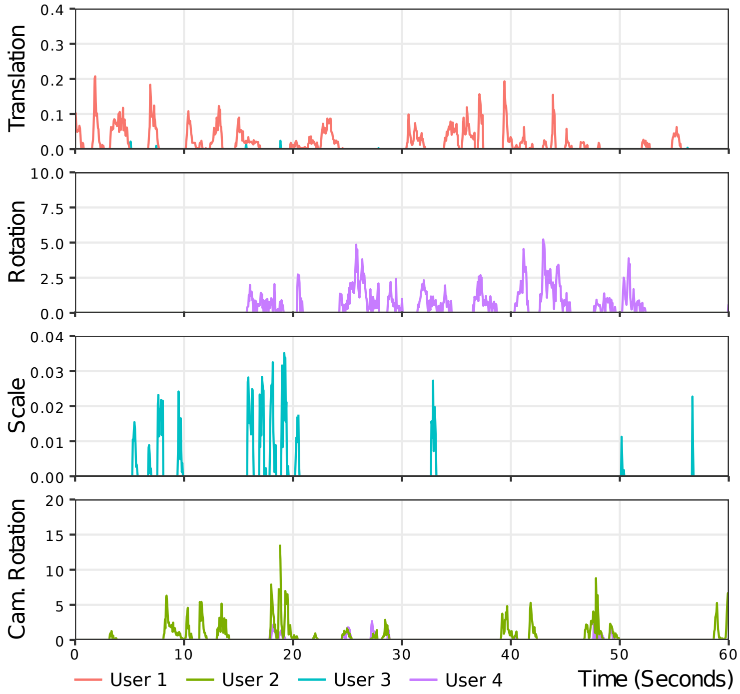
3.4.2 Groups vs. Work Division

3.4.2.1 Accuracy vs. Work Distribution Balance

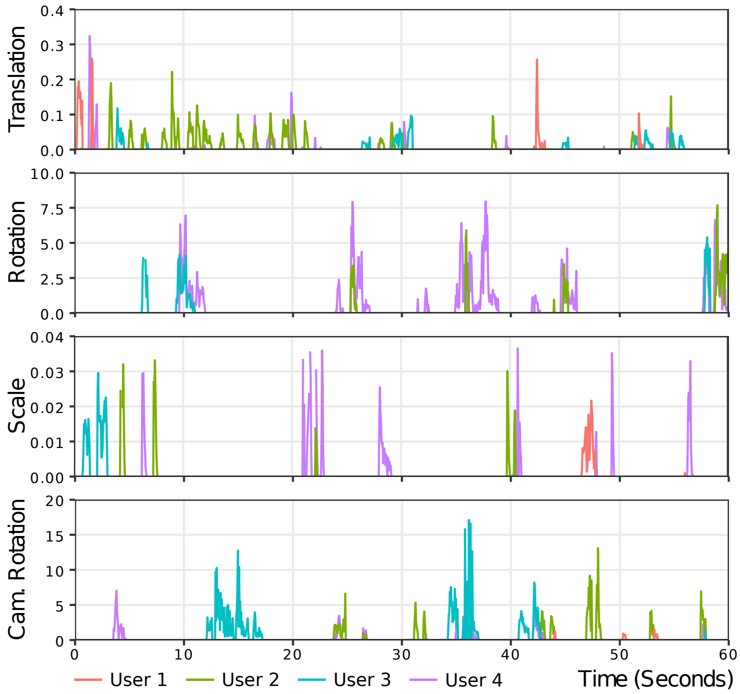
For a more complete understanding on how the various group sizes affect accuracy, we analyzed the balance of the work distribution among members of teams with different sizes. Our hypothesis is that the balance of work distribution affects the task accuracy. The work distribution balance is a value between zero and one, where zero is the minimum work distribution and one is the maximum work distribution.

Before the work distribution evaluation, we performed an analysis to assess the workload of each team members for all group sizes. Each participant's workload was quantitatively measured dividing the user's active time by the group's active time. The

Figure 3.7: Time series of one task showing the collaboration strategy adopted by Team A and B with four participants. Both teams have similar work division balance (A = 96.4%, B = 95.5%) but different strategies. Team A splitted the transformations among the participants in such a way that, on average, all team members performed 3.25 roles swap. They completed the task in 60.7 seconds and with 1.85 errors. Team B did not divide the transformations, on average all team members performed 15 roles swap. They completed the task in 63.9 seconds (the last 3.9s are omitted in the plot) and with 2.78 errors.



(a) Team A



(b) Team B

Source: the Author

Kruskal-Wallis test revealed a significant effect in the workload when group sizes vary (H(3) = 79.0784, $p < 0.0001$). The Dunn post-hoc indicates significant decrease in workload between groups size 1 and 2 ($p < 0.0006$), 1 and 3 ($p < 0.0001$), 1 and 4 (p = 0.0031), 2 and 3 ($p < 0.0007$), 2 and 4 ($p < 0.0001$) and 3 and 4 (p = 0.0303).

Then, we computed the work distribution balance. We calculated the group variance using each individual workload previously calculated. We adjusted the results by multiplying the variance by the respective group size. In this test, only groups with two, three and four members were evaluated, since the work distribution in groups with one participant is zero. However, the Kruskal-Wallis test failed to reject equality of medians across different group sizes for work distribution balance (H(2) = 4.482, p = 0.11). This hinders our intent of correlating accuracy with overall work distribution balance directly. Nevertheless, this result permitted us to focus on action division patterns, as detailed next.

3.4.2.2 Groups vs. User Roles

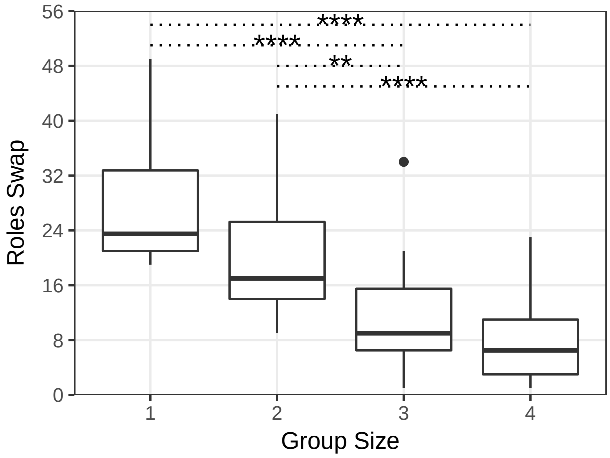
The work distribution balance omits the group action division by calculating a single distribution score for the whole group. Here, we analyze each team member individually to find a work division pattern. Our hypothesis is that the groups that divide actions among members are more accurate than the groups that do not divide. The time series on Figure 3.7 shows the different strategies adopted by two groups with four participants, each with similar work division balance (A = 96.4%, B = 95.5%) but with different action strategies and accuracy. To extract the division of actions between the groups, we identified the frequency users change between actions (translation, rotation, scale). We call this change a *role swap*.

The Kruskal-Wallis variance analysis revealed significant effect between groups and user roles changes (H(3) = 40.1615, $p < 0.0001$). The post-hoc Dunn test indicates that significant action division occurs between groups size 1 and 3 ($p < 0.0001$), 1 and 4 ($p < 0.0001$), 2 and 3 (p = 0.0031) and 2 and 4 ($p < 0.0001$). Groups size 1 and 2, 3 and 4 do not differ significantly (see Figure 3.8).

3.4.2.3 Accuracy vs. User Roles

Since the accuracy and the action division have similar behaviors, we hypothesized that the two variables were related. The Pearson correlation revealed a significant effect between error and user roles changes (r = -0.3957, $p < 0.0001$).

Figure 3.8: Division of actions between the groups. Individuals in groups with four members swap roles less often (Mdn = 6.5, SD = 6.26), followed by groups with size 3 (Mdn = 9, SD = 7.21), size 2 (Mdn = 17, SD = 10.22) and size 1 (Mdn = 23.5, SD = 10.71).

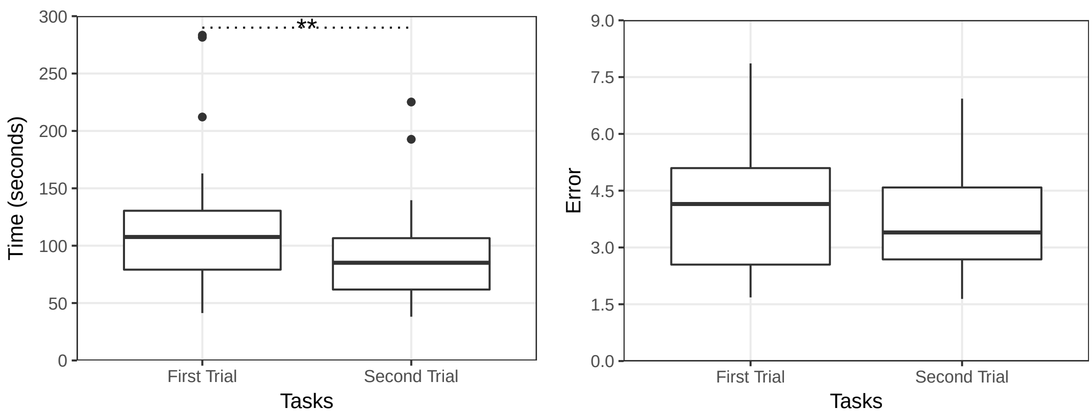


Source: the Author

3.4.3 Learning Between Trials

In Figure 3.9, we measured the learning effect between the two trials. We assessed the time and accuracy for all groups together. The Wilcoxon signed rank test indicated no significant increase in accuracy between trials (Z = 144, $p < 0.5879$) and a significant effect in time between trials (Z = 214, $p = 0.0032$).

Figure 3.9: Learning Between Trials. (a) Groups complete the task significantly faster in the second trial (Mdn = 85.12, SD = 44.81) than the first (Mdn = 107.59 , SD = 65.58). (b) There is no significant accuracy effect between the first (Mdn = 4.14, SD = 1.74) and the second (Mdn = 3.59, SD = 1.83) trials.



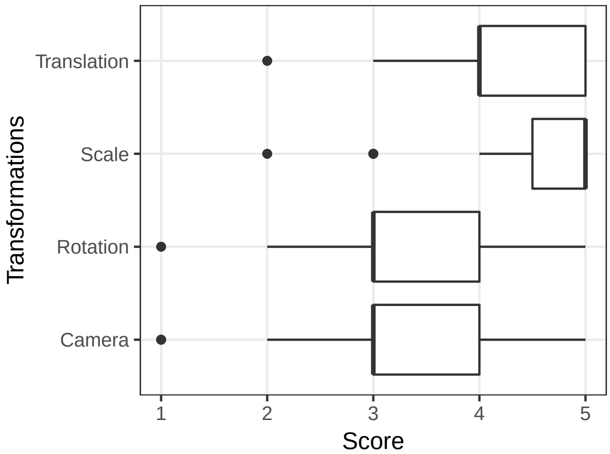
(a) (b)

Source: the Author

3.4.4 User Comfort

In Figure 3.10, we report the user responses to comfort when performing the transformations. Most users reported that the translation and scale actions are very comfortable to perform. Object and camera rotations had mostly neutral or positive opinions. They have almost the same comfort level according to users, being slightly but clearly below the other two transformations.

Figure 3.10: User's opinion about the technique comfort for each transformation. The results are reported in the Likert scale where 1 is very uncomfortable and 5 is very comfortable. Translation (Mdn = 4, SD = 0.73), scale (Mdn = 5, SD = 0.65), rotation (Mdn = 3, SD = 0.94) and camera rotation (Mdn = 3, SD = 1.08).



Source: the Author

3.5 Discussion

3.5.1 Technique Performance

The findings in our study indicate that handheld devices are powerful tools for 3D collaborative tasks. We observed that team arrangements with two, three and four members solve the tasks significantly more accurately than individuals (groups with one member). Interestingly, the accuracy increases faster with larger group sizes. For group sizes of one and two, there is only significant increase when compared with groups with two or more additional members. However, between group sizes 3 and 4, the accuracy increases significantly. No significant difference in completion time is reported, which may have been influenced by the instruction to prioritize accuracy over time in our experiment.

To understand the causes for accuracy increasing with larger group sizes, we investigated whether it was caused by the way teams balance the actions among team members. The statistical analysis revealed no significant increase. Thus, we further explored the accuracy effect by analyzing the team's work division. We observed a significant drop in the swap of roles of the team members on larger groups. It indicates that users tend to specialize in one transformation, consequently better dividing the tasks. The behavior between accuracy and work division follows the same trend, as observed in Figure 3.6b and Figure 3.8. We correlated the two variables, and the results indicate that the work division strategy is related to the error drop.

We also investigated the learning effect between the two repeated trials. We report a significant drop in time to complete the tasks without affecting accuracy. It indicates that with more training, groups tend to perform the task faster while keeping equivalent accuracy.

3.5.2 Technique Comfort

The results show that the technique is comfortable. We observed that translation and scale received better scores than both object and camera rotations. It is known that actions performed by small muscle groups, such as fingers tend to be more comfortable than larger muscle groups, such as wrist and arm (ZHAI; MILGRAM; BUXTON, 1996). This may explain why the rotation actions were ranked one step below the others, even if still well ranked. Nevertheless, we exchanged comfort for affordance in our design. We believe that the use of direct mapping for rotations is more natural than rotations with the fingers, as the user indirectly holds the virtual object.

3.5.3 Ringelmann Effect

The Ringelmann Effect states that the addition of new co-workers in a collaborative task (the original was a rope pulling task) leads to a linear decrement in the member's performance. Ingham et al. (1974) reproduced the effect with teams of 1 to 6 participants. They report a significant drop in individual performance in groups with more members. Even though our experiment was not designed to test the effect, in Section 3.4.2.1, we found a correlation between workload and group size. It suggests that, in virtual collaboration, team members tend to work less in larger groups as previously demonstrated in the real world collaborative tasks.

3.6 Chapter Summary

We presented the design of a novel 3D user interface for collaborative object manipulation in 3D virtual environments in order to evaluate the cooperation in groups. The technique is based on smartphones and uses the touchscreen and the inertial sensors as a 3DUI. The technique has proven to be intuitive and robust. More than a hundred participants have tested it and could provide data and feedback about its usability. In several informal tests, users were invited to download the app in their own smartphones and to join an on-going manipulation session. We tested it with teams composed of more than ten members using smartphones of different models. All of them easily and quickly understood the purpose and mechanics of the technique, indicating a high affordance.

In formal user experiments, we demonstrated that teams with more members performed more accurately than smaller groups. In larger groups, members could better divide the basic tasks, assuming roles and, as a consequence, decreasing the workload. Speed was not the goal in our experiments, but we also observed that speed increases from the first use of the system to the second one without affecting accuracy, demonstrating that the users learn fast. However, more tests should be done regarding the technique learning curve. Our findings indicate that in virtual manipulation tasks, where precision is mandatory, sharing the manipulation aspects significantly increases the accuracy.

Smartphones, being versatile and ubiquitous devices, conveniently adapt as input devices to be used in combination with an external or wearable display. While we evaluated our interaction technique using a large screen (integrating a shared+personal view) as display device, the technique is friendly to other display form factors such as HMDs or CAVEs. Changing the view to VR HMDs or augmented reality could provide greater individual perspective, while the input manipulations would not change much. Those aspects are further explored in Chapters 4 and 5.

We reported the design, assessment and results presented in this Chapter in two papers. The first paper presented the initial interface design and was published as part of the IEEE 3D User Interfaces Contest in 2016 (GRANDI et al., 2016). Our solution won the first prize among eight other contenders. Then, we published the collaborative evaluation results in the ACM CHI Conference on Human Factors in Computing Systems in 2017 (GRANDI et al., 2017b).

4 COLLABORATIVE 3D INTERACTION TECHNIQUE FOR HANDHELD AUGMENTED REALITY

The premise of augmented reality (AR) is to enhance sensory perception through computer-generated information, mainly visual information, from virtual objects to metadata about the environment. Different AR displays exist, such as head-mounted see-through displays, surface mapped projections and video-mediated rendering. The latter includes handheld devices containing a rear camera and a screen, such as mobile phones and tablets.

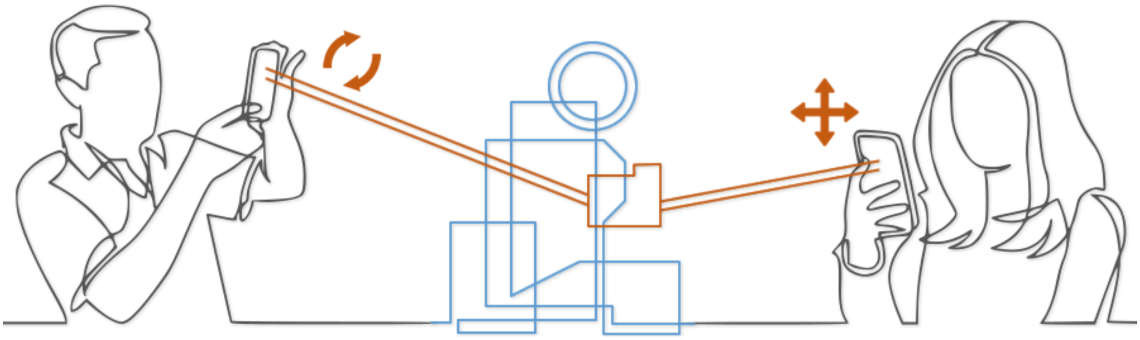
Current handheld devices seem to be the ideal device to fulfill AR requirements for everyday applications. They have sensors to capture touch, movement, and image. This allows the use of computer vision and integration of inertial sensors to define the pose of the device in the physical space. Virtual elements then overlay the real world captured by the device's embedded camera directly on the mobile device screen. Touchscreen gestures are widespread among users and are being smoothly adopted to trigger interactive content. Besides, interest for mobile AR increased with the appearance of games such as Pokemon GO, reaching a broad audience. Very recently, Apple and Google released their APIs (ARKit [1] and ARCore [2]) for native AR support in their operational systems.

While some virtual augmentations are informative, others are interactive. Interactive content may be in various formats, but the most common are three-dimensional objects. The manipulation of 3D objects in AR environments is a complex task that requires the control of multiple degrees-of-freedom (DoF) for selecting, translating, rotating and scaling objects. Moreover, virtual objects are intangible, and interaction with them can only be achieved through full-body tracking or mediated by a handheld device. Touch gestures on a 2D screen provide a straightforward method to interact with virtual objects. Alternatively, 3D interactions with the device movements (SAMINI; PALMERIUS, 2016) and around-the-device (KIM; LEE, 2016), provide a direct mapping between the input and the respective manipulation. Furthermore, combined with touch gestures these interactions provide intuitive, high precision and fast control in multiple DOF.

The widespread availability of mobile phones allows users to access AR through a personal perspective, and to share and collaborate with other users when interacting with virtual content (BORING et al., 2010; MÜLLER; RÄDLE; REITERER, 2016). Since it is interesting that multiple people cooperate in virtual spaces, an interface capable of handling

[1] https://developer.apple.com/arkit
[2] https://developers.google.com/ar

Figure 4.1: Two users simultaneously manipulating a virtual object in augmented reality. Each user can have a different perspective view of the scene. Rays are drawn from the device to the selected object to inform users the current selection. Virtual icons indicates the current transformation.



Source: the Author

and synchronizing inputs of many users for cooperative work is desirable. However, as of today, research in the 3D object manipulation field mainly focus on single-user interaction.

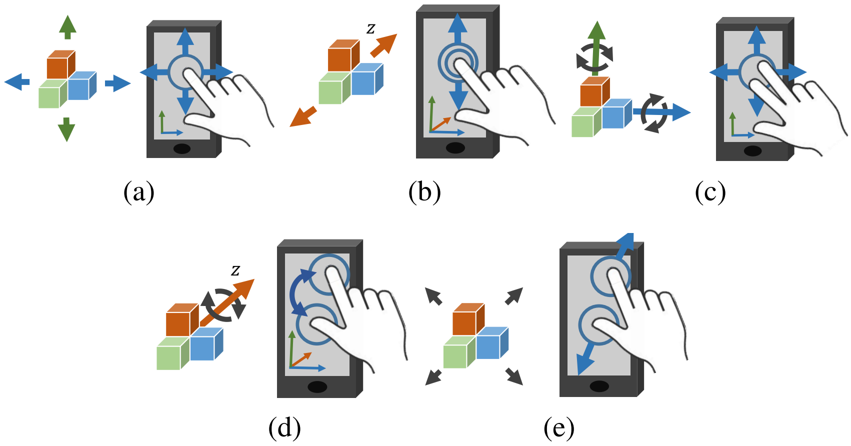
In this chapter, we present a novel collaborative 3D user interface for virtual objects' manipulation in handheld augmented reality (Figure **??**). Our solution creates a shared medium where several users can interact through their points-of-view and simultaneously manipulate 3D virtual augmentations. We integrate touch gestures and device movements into a hybrid manipulation for fast and precise interaction. Our technique handles inputs from several participants with their devices. Participants can either perform manipulations alone or manipulate objects together, simultaneously. We implement UI elements to keep users aware of the other's actions. The design to combine all these features makes our approach unique and is the main contribution of the paper.

Besides, we present two experiments. In the first, we evaluated the interaction interface where we compare the single user performance in three interaction conditions: touch gestures, movements and hybrid. Then, we conducted a study to understand and classify the strategies that arise while working in pairs, when partners are free to make their task organization. Furthermore, we investigated the effectiveness of simultaneous manipulations compared with single user manipulations.

4.1 Design of a Collaborative Handheld AR Technique

We propose a flexible set of actions for manipulation using handheld-based devices. It engenders the availability of *touch gesture* and *device movement* inputs that aid both precise and fast spatial transformations in augmented reality scenarios. The user decides

Figure 4.2: Manipulations performed over the selected object: (a) A touch and slide translates in xy axis, (b) one tap followed by a touch and slide translates in z axis, (c) touch and slide with two fingers rotates in xy axis, (d) touch and rotate two fingers rotates in z axis and (e) pinch and spread of two fingers to uniformly modify the object's scale.



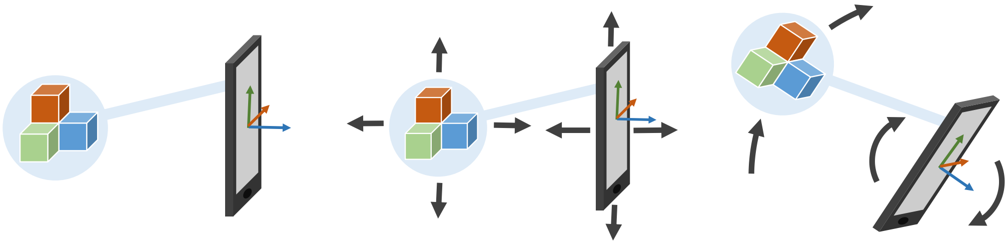
Source: the Author

the most appropriate input depending on the task needs. A simple touch+hold action on the screen surface switches between touch gesture manipulation (Sec. 4.1.1) and device movement manipulation (Sec. 4.1.2). Moreover, our technique was designed to support an unlimited number of simultaneous participants. All users have access to all available functions and can apply transformations to different objects or to the same object simultaneously while observing the scene from a different point of views.

4.1.1 Touch Gestures Manipulations

We convert finger gestures on the touchscreen into 3D transformations to manipulate a total of 7 DOF of a selected 3D object. More specifically, there are 3 DOF for translation, 3 DOF for rotation and 1 DOF for uniform scale. The transformations are applied relative to the touchscreen plane orientation (i.e., the device orientation), similarly to the proposed by Grandi et al. (GRANDI et al., 2017b) and Katzakis et al. (KATZAKIS et al., 2015) (Figure 4.2). The transformations are performed using one and two fingers. We based our touch gestures implementation on the *DS3* technique (MARTINET; CASIEZ; GRISONI, 2012) with variations. The touch and slide of one finger in the device xy orientation plane move the object in the same direction as the finger slide. The gesture sequence of one tap followed by a touch and slide of one finger horizontally moves the object towards the z device axis, unlike the *DS3* that uses another finger for z translation (Figure 4.2a-b). Two fingers touch and slide enables rotation. The slide of two fingers in a horizontal direction

Figure 4.3: Manipulations performed on the attached object: (a) object-device attachment, the object and the device keep an invariant rigid transformation. In this way, the object translates (b) and rotates (c) with the device movements.



(a) Object attached (b) Object translation (c) Object rotation

Source: the Author

affects yaw, vertical sliding changes pitch and pivoting affects roll rotations (Figure 4.2c-d). Finally, we add the pinch and spread of two fingers to uniformly modify the object's scale regardless of the screen plane orientation (Figure 4.2e). The rotation and scale gestures are not combined. Thus, it is necessary to release the fingers and start another gesture to change the transformation. All transformation modes are correctly applied depending on the gesture. Therefore, there is no necessity for UI buttons to switch modes.

4.1.2 Device Movements Manipulations

This approach attaches the object to the physical pose of the device (SAMINI; PALMERIUS, 2016). The transformation with movements is activated by pressing and holding a circular button in the lower-right corner of the interface and is halted once the button is released. While a finger is pressing the button, the object translates and rotates with the device while keeping an invariant rigid transformation relative to it (Figure 4.3). We do not allow changes in the transform rate and transformations decouple to preserve the absolute mapping. Clutch can be used to reach a total rotation beyond arms limits (i.e., perform object rotation, reposition the device, then perform a new object rotation). Depending on the object-device bound distance, the rotation affects the object position, as shown in Figure 4.3c. While moving the object, it is possible to slide another finger on the screen closer or far away to scale the object.

Equation 4.1 defines the matrix calculation to transform the object during the manipulation.

$$T = V(V^{-1}V_{prev})V^{-1} \tag{4.1}$$

The transformation first converts the model to the camera coordinate system, symbolized by the matrix V. After, we apply to V the same transformation that the camera suffered during the last iteration, for that, the view matrix of the previous frame is used, represented by V_{prev}. Finally, the transformation is multiplied by the inverse of the view matrix to return the model to the world coordinate system.

4.1.3 Selection and Grouping

Selection is performed by touching the desired virtual object on the touchscreen. A single tap selects only one object. It is also possible to select more objects at the same time. Having selected an object, the user can touch and hold on a second object to start the multiple selection mode. Once in this mode, single touches select or deselect objects. A touch on an empty area deselect all objects.

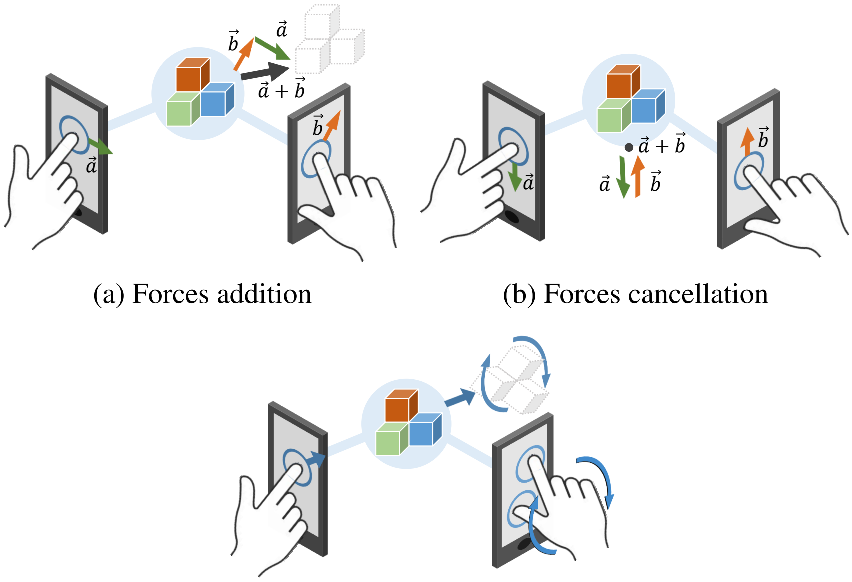
Grouping is used to save a multiple selection. The user can make a group by pressing the *group* button in the lower left corner of the screen. This button only appears if more than one object is currently selected. A selected group can be disassembled by pressing the same button, which now contains the text *ungroup*. A group obeys the same interaction rules as a single object. So, it is possible to select multiple groups and merge them into a new group and edit a group by touching and holding an object to add or remove them from the group.

4.1.4 Simultaneous Manipulations

Our technique supports simultaneous manipulation of virtual objects by multiple users implementing the *Sum-of-Contributions* model (Chapter 2.4.1). While working in groups, users can either independently manipulate objects or interact simultaneously with the same object. Thus, cooperative strategies can be established depending on the task needs and the ability of the team members.

When two or more users are manipulating the same object, the actions performed by each individual counts as a transformation step. We multiply each user transform matrix by the virtual object transform matrix. Thus, every contribution from each user is summed up in the final object's transformation without restrictions or weights (Figure 4.4a). Therefore, if two users move the object in opposite directions, the position of the object will not

Figure 4.4: Concurrent access to transformations. Every action performed by each user counts as a transformation step. In (a) and (b) the final translation is the sum of the transform vectors. If users apply a translation in opposite directions, the object will not move. If they manipulate different transformations in parallel, the transformations are merged (c).



(a) Forces addition

(b) Forces cancellation

(c) Transformations merge

Source: the Author

change (Figure 4.4b). On the other hand, if they manipulate the different transformations in parallel, the transformations are combined (Figure 4.4c). The simultaneous manipulations can occur with any combination of *touch gestures* and *device movements*.

We added two virtual elements to make users aware of the other participants' actions in the virtual scene. Regarding selection, we draw rays from the device location to the currently selected objects. The ray informs about the focus of interest of other users interacting in the same AR environment, as shown in Figure A.2. We distinguish between colors the user selection ray and the other participants' selection rays. We also render icons on the virtual rays indicating the transformation being performed by each user, so that users can be aware of each other's actions without the need for verbal communication. The icons can indicate that a mobile phone is either performing a movement or translation, rotation or scale with touch on the selected object.

4.2 Public Demonstrations

We conducted two public demonstrations of our collaborative AR interface prior to formal assessments. In these demonstrations we presented a constructive task where multiple users could interact with 3D shapes in order to assemble structures (see Figure 4.5). We did not impose any restrictions on the creation. Over fifty people have tested our technique in the demonstrations. We registered periods in which over five people were interacting at the same time. We collected users' feedback to refine the collaborative interaction and experience.

Figure 4.5: The setup presented in the two public demonstrations. Multiple users could interact with 3D shapes and build structures.

Source: the Author

4.3 Experiment 1: Single User Assessment

Our goal is to compare task completion time and error rate between the three 3D manipulation methods: *Touch gestures*, *Device movements* and *Hybrid*. Thus, we carried this experiment with a single user. We hypothesize the *Hybrid* technique to be the fastest and will have the lowest error rate. Theoretically, the *Device movements* technique has a time advantage over the *Touch gestures* since the manipulation is analogous to a person carrying an object. However, the physical effort required by the *Device movements* technique may introduce more error during precise positioning, which is less expected with the *Touch gestures* technique.

4.3.1 Task and Stimuli

We designed a 3D docking task that comprises translation in 3-DOF and orientation in 3-DOF. A docking task consists of transforming a virtual object to a target position and orientation, and it is a widely adopted task to evaluate interfaces and techniques for spatial manipulations (FROEHLICH et al., 2006; VUIBERT; STUERZLINGER; COOPERSTOCK, 2015). In our experiment, we asked participants to dock a virtual *moving piece*, controlled by the user with a similar virtual *static piece*. Both *moving piece* and *static piece* had the color matched. The *static piece* was 50% semi-transparent while the *moving piece* was opaque (Figure 4.12).

The pieces configuration stimuli are composed of cube blocks similar to the Shepard and Metzler (SHEPARD; METZLER, 1971) construction. The blocks are assigned with different colors to avoid ambiguity when matching the target piece (Figure 4.12). The blocks have 6cm long edges.

For each docking task, only one *moving piece* and their respective *static piece* appears in the scene. We placed the *static piece* always in the center of the scene while the *moving piece* has four possible spawn positions. The distance between the two pieces at the start of a trial is fixed at 35cm (Figure 4.6).

The task's mechanics summarized:

- Aim the mobile device toward the table.
- Tap on the *moving piece* to select it.
- Position the *moving piece* on the *static piece* with the available inputs
- Confirm the docking with the button on the lower left corner of the device.

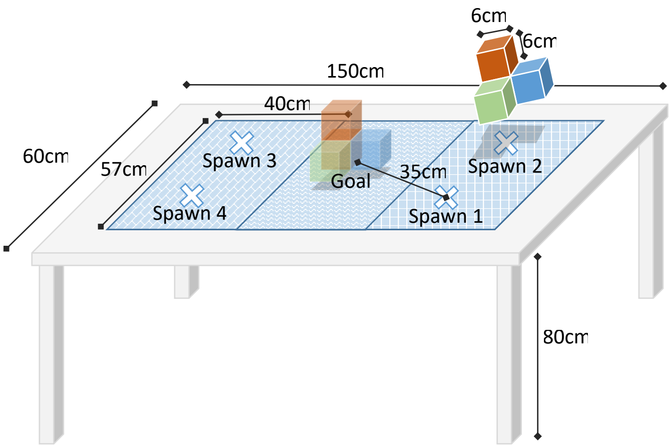
We render shadows and apply collision forces between the piece and the table to provide additional depth cues and interaction with the physical environment.

4.3.2 Apparatus

The experimental setup is composed of an Apple iPad mini 4 with 7.9 inches screen (approx. 324 ppi pixel density) and weights 299g. The 8MP rear camera and the Vuforia SDK [3] were used to track the mobile device physical pose relative to the fiducial markers.

[3] www.vuforia.com

Figure 4.6: The setup of the experiment. Four spawn point with the same distance to the goal. The working area is 150x60cm. Each block that composes a piece has 6cm long edges.



Source: the Author

We used multiple markers and the *extended tracking* feature to extend the interaction range. These markers were placed on top of a table. We used a server application to manage and send to the mobile device the experiment parameters (modality order, trials randomization, and piece spawn position) and to record all user interactions during the experiment, a dedicated WiFi connection and a table with patterns for tracking. Both server and client's application were developed using the Unity3D game engine [4] and the communication is made with the Unity UNET network API.

4.3.3 Subjects

Twenty subjects participated voluntarily in this experiment (six female), aged 25.05 years in average (SD=3.27). All subjects read and agreed to an Informed Consent Form (see Appx. C) before the experiment. They were all Computer Science students. Two of them reported minimal movement restrictions on the wrist and finger that did not affect their performance during the experiment. They all had either normal or corrected to normal vision.

[4] www.unity3d.com

4.3.4 Experimental Setup

The experiment follows a repeated measures within-subject design with *Technique (Touch gestures, Device movements and Hybrid)* and *rotational angles (45° and 90°)* as the independent variables. The dependent variables were *time* and *accuracy* (position and orientation) to complete each docking. We have also collected the user *transformation actions (translation, rotation)* and the *user's physical positions* in the environment for post hoc association with the aforementioned dependent and independent variables.

The participants answered a characterization form and performed a mental rotation test at least one day before the experiment. The mental rotation experiment used is similar to that used by Shepard and Metzler (SHEPARD; METZLER, 1971). The test was used to assess the participant's ability to understand 3D rotations. In the experiment session, before the trials,the participants were guided to experiment the input commands. In this phase, only one object was displayed and the participant had no target objective.

The presentation order of the three conditions was counter-balanced that resulted in 6 different group orders. For each input technique, participants had one practice session composed of two docking trials and a recorded session composed of eight docking trials.

In the first practice trial, we displayed, in real-time on the device's screen, the actual values of position and orientation errors. The text values changed colors (from white to green) for each parameter to inform when a threshold (1.5cm and 15° difference (VUIBERT; STUERZLINGER; COOPERSTOCK, 2015)) was achieved. For the second practice trial, the reference position and orientation errors were displayed after the participant confirm and finish the docking. Then, participants were asked to perform the eight valid trials. We asked participants to balance accuracy and speed. No reference errors were displayed during the recorded trials to avoid a bias toward accuracy (HINCKLEY et al., 1997). Participants were orally informed of their progress when four and two trials were missing to complete the block of the trials. Each trial started when the user selected the virtual object and finishes when the user confirms the docking by pressing a button in the lower left corner of the device's screen. Each virtual piece appears in sequence after the previous docking confirmation. It was possible to select only the *moving piece* and once selected it could not be unselected. After the recorded session, participants were allowed to rest and were asked to assess their workload level with the NASA's Task Load Index [5].

In summary, the experiment consisted of: **20** participants × **3** techniques × **8** trials

[5]https://humansystems.arc.nasa.gov/groups/tlx/downloads/TLXScale.pdf

(4 - 45° and 4 - 90° of rotational difference) = **480** unique docking.

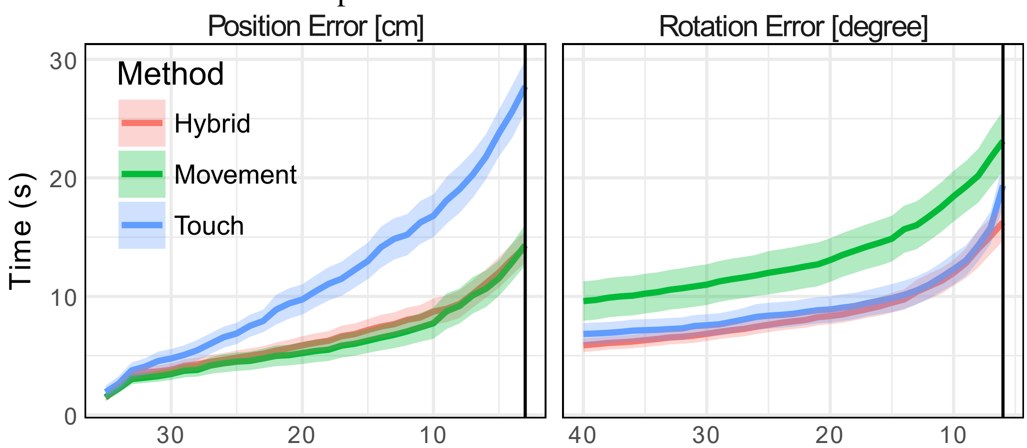
At the end of the experiment participants answered the Single Easy Question (SEQ) (SAURO; DUMAS, 2009) for each manipulation condition ("Overall, How difficult was the tasks with *condition*?"). The SEQ was rated on a 7-point Likert scale, ranging from 1 (“Very Hard”) to 7 (“Very Easy”). Then, we applied a System Usability Score (SUS) questionnaire to assess the overall usability of the experimental setup (see Appx. E). The experiment sessions lasted approximately 50 minutes.

4.4 Results

4.4.1 Accuracy and Time

We removed 3 trials where subjects failed to attain the minimal precision of 5 cm and 15° relative to the reference docking object. For the statistical analysis, we take the median of the time, translation error and rotation error for each combination of method and rotation angle per subject. Repeated measures ANOVA was used to test the statistical significance of the manipulated factors. For the precision analysis, we consider the precision attained by the subjects when they indicated that they were satisfied with the docking.

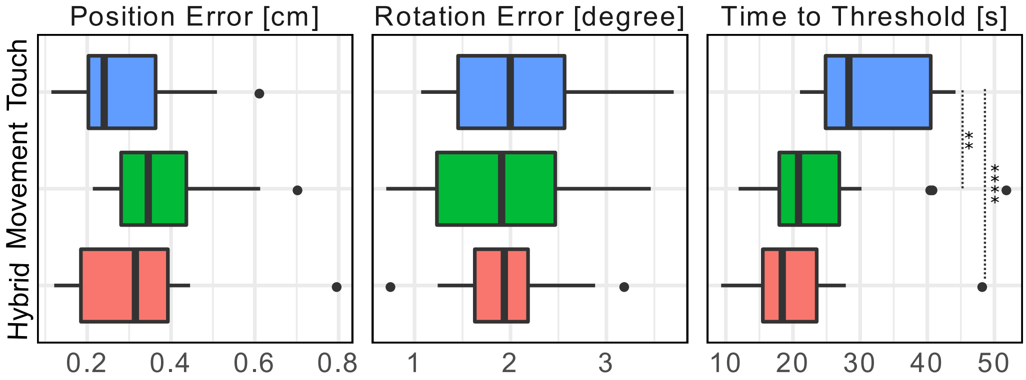
Figure 4.7: Error reduction by time for position and rotation. The faster the error is reduced, the better. The shaded areas represent the standard error of the mean.



Source: the Author

Figure 4.7 shows the time by error reduction rate for position and rotation. The smaller amount of time needed to reduce the error suggests that *Movement* manipulation is more efficient for positioning the object, while *Touch* manipulation is more efficient for rotating the object. It also suggests that subjects could take advantage of the *Hybrid*

Figure 4.8: We found a statistically significant advantage of Hybrid over Movement and Touch, and of Movement over Touch for the time using a 1.15cm and 8° threshold. There is no significant difference in position error and rotation error between interaction methods.



Source: the Author

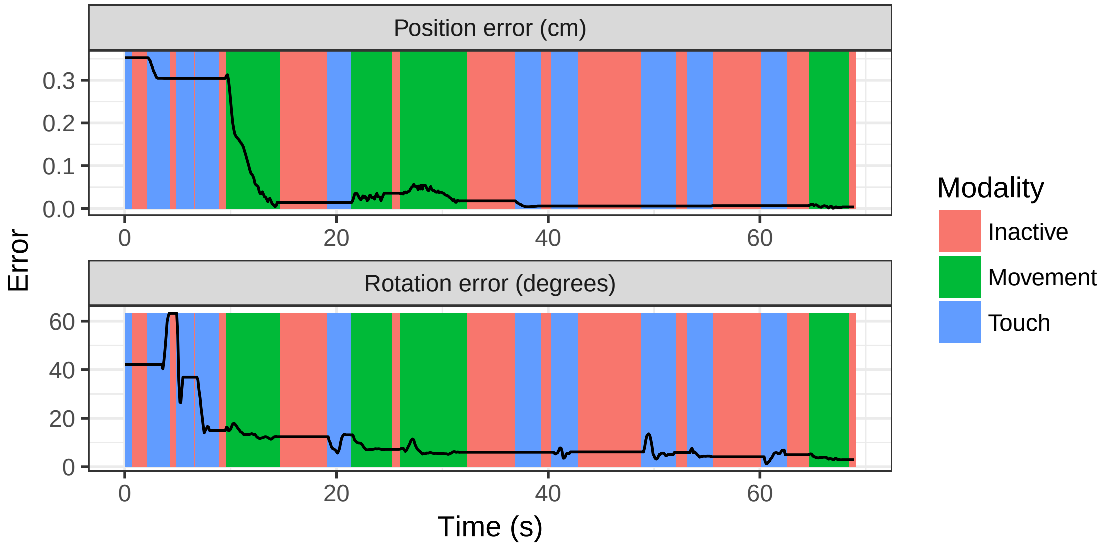
approach, presenting an error reduction rate very similar to *Movements* regarding positioning, and *Touch* in term of rotation. We test the statistical significance for the point in time when subjects achieved a minimal precision of 1.15cm and 8°, these values were chosen as they represent the performance attained in every trial of the experiment. Figure 4.8 shows the time to reach the threshold for each method and the position and rotation errors achieved when users were satisfied with the docking. We found a statistically significant difference for input method ($F_{(2,38)} = 24.9$, $p < .001$) but not for the initial rotation factor ($F_{(1,19)} = .5$, $p > .48$), nor their interaction ($F_{(2,38)} = .03$, $p > .96$). Post-hoc T-test of the method indicates that *Hybrid* was more efficient than *Movement* ($t_{(19)} = 3.5$, $p < .005$) and *Touch* ($t_{(19)} = 6.7$, $p < .001$) manipulation, and that *Movement* had advantage over *Touch* ($t_{(19)} = 4.2$, $p < .001$). Our results point that subjects effectively took less time when interacting with the hybrid approach.

On the other hand, we found no significant effect of *Technique* on position and rotation error. That is, subjects could dock the objects with similar precision regardless of the *Technique* in use.

4.4.2 User Docking Behavior

Figure 4.9 shows the time series of one trial using the hybrid method. The black line represents the adjustments that the user made to position and orientate the virtual object during the task. The background colors represents the modality (*device movements, touch or inactive*) used in each time stamp.

Figure 4.9: Time series of one trial using the hybrid method. The black line is the error for position and rotation and the colors represents the modality used during the task.



Source: the Author

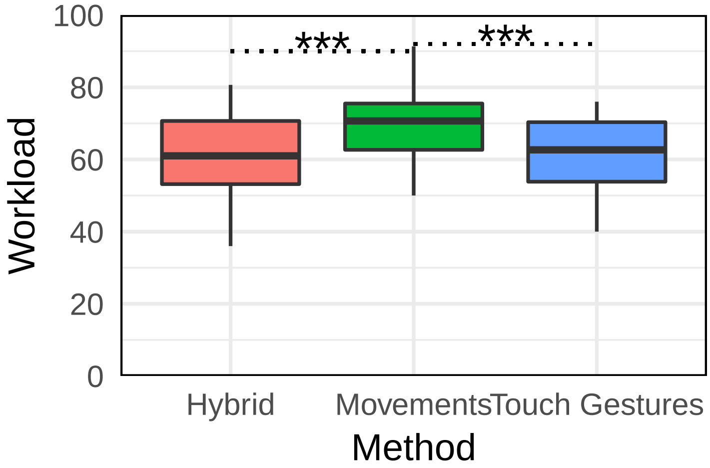
4.4.3 Workload, Usability and Single Easy Question

Non-parametric Friedman tests were conducted to compare the effect of the three manipulation techniques (*Hybrid, Movements and Touch Gestures*) to the NASA TLX workload questionnaire (Appx. F) and Single Easy Question (SEQ) (Appx. D). The post-hoc analysis was conducted with Wilcoxon signed-ranks test with a Holm-Bonferroni correction for multiple comparisons.

Figure 4.10 shows the result of the NASA TLX user workload for each tested condition. The result of workload for each tested condition revealed a significant effect between conditions ($X^2(2) = 14.354$, $p < .001$) with *Hybrid* $Med = 61$,$IQR = 17.5$, *Movements* $Med = 70.7$, $IQR = 12.8$ and *Touch Gestures* $Med = 62.7$, $IQR = 16.5$. The post-hoc test indicates that significant workload increase occurs from *Hybrid* to *Movements* ($p < .006$) and from *Touch Gestures* to *Movements* ($p < .007$). No significant workload differences occurs between *Touch Gestures* and *Hybrid* ($p > .7$).

An analysis of each individual NASA TLX factor revealed a significant effect on Performance ($X^2(2) = 6.51$, $p < .04$), Effort ($X^2(2) = 7.42$, $p < .03$) and Frustration ($X^2(2) = 14$, $p < .001$). Significant Effort factor increase occurs from *Hybrid* to *Movement* ($p < .025$) and from *Touch Gestures* to *Movement* ($p < .02$). Significant Frustration factor increase occurs from *Hybrid* to *Movement* ($p < .003$) and from *Touch Gestures* to *Movement* ($p < .046$). No significant difference was found for Performance on the Post-hoc analysis.

Figure 4.10: Weighted NASA TLX scores for each factor and workload for Hybrid (M = 60.90, SD = 11.69), Movements (M = 68.93, SD = 10.77) and Touch Gestures (M = 62.10, SD = 10.07) conditions.



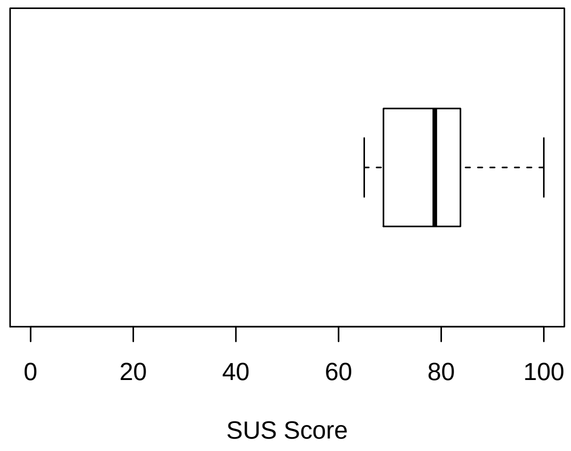
We asked participants a Single Easy Question (SEQ) (SAURO; DUMAS, 2009) to compare the three conditions ("Overall, How difficult was the tasks with <condition>? "). The SEQ was rated on a 7-point Likert scale, ranging from 1 ("Very Hard") to 7 ("Very Easy"). We found a significant effect of manipulation condition ($X^2(2) = 27.757$, $p < .001$). Post-hoc indicates a significant difficulty increases from *Hybrid* to *Movements* ($p < .001$), from *Hybrid* to *Touch Gestures* ($p < .001$) and from *Movements* to *Touch Gestures* ($p < .001$). The *Hybrid* was ranked as the easiest ($Med = 6$ $IQR = .25$) and the *Movements* as the hardest ($Med = 3$ $IQR = 1$), while *Touch Gestures* ($Med = 5$ $IQR = 1$) was ranked between two other conditions.

Figure 4.11 shows the System Usability Score (SUS). The mean scores ranged from 65 to 100 (M=77.12, SD=9.5). According to surveys that compare SUS scores for different systems, the system is ranked as a "Good" (BANGOR; KORTUM; MILLER, 2009) (see Appx. E).

4.5 Experiment 2: Pair Work Assessment

The pair work assessment described here focuses on the evaluation of collaborative aspects when two users are manipulating virtual objects in the same scene. Different from works that impose and compare different collaborative strategies (LIU et al., 2016), we aim at observing and classifying the strategies that emerge when users are free to make their task organization. During public demonstrations of our collaborative AR interface, we observed that groups often adopted different strategies to accomplish a constructive task.

Figure 4.11: Mean SUS score for all subjects (M=77.12, SD=9.5).



Source: the Author

Thus, in this experiment, we propose to control the level of occlusion in the augmented scene to observe how interaction strategies change, and how these strategies compare regarding performance.

We expect that users will tend to organize themselves depending on the level of occlusion in the augmented scene. We hypothesize that two strategies will appear: *Independent Interaction*, where users divide the problem and each solves part of it as in a single user approach, and a *Shared Interaction*, where the task is performed sequentially with both users focusing attention on the same sub-task. Moreover, we would like to analyze if trials where users apply the *shared interaction* approach will have any effect on time when compared with trials performed separately as in single user mode.

4.5.1 Task and Stimuli

The virtual scene setup of the previous experiment was reused for this experiment. We added collision detection between pieces and gravity attraction besides the already included shadows to make the pieces act similarly as physical blocks. Differently from the previous experiment, where only one pre-selected *movable piece* and one *static piece* were shown at a time, here, all *movable pieces* are shown and the users need to select and dock the correct piece. Each *static piece* property – spawn position, rotation and scale – is randomly chosen from a list of valid transformation. Since scale is introduced in this experiment, we defined four possible scales: 60%, 80%, 120%, 140% of the *movable* *piece* size. When one *movable piece* is docked, both the *movable piece* and the respective *static piece* disappear and a new *static piece* appear in the next docking space. The task is completed after all pieces are docked.

To stimulate cooperative work, we created three conditions where we vary the level of occlusion in the scene. One with no occlusions, the second with moderate occlusion, and the last condition with high occlusion. For that, we added virtual walls on the augmented scene to force users to look for new vantage points from which the objects are not occluded (Fig. 4.12). The working space had 57x120cm and was divided into a 57x40cm docking space, where the *demands* appear, and the supplies space (57x80cm), where the 18 pieces are initially placed. The walls are 24cm high (the exact height of two stacked blocks) and 2cm thick. The longer walls have 80cm and the shorter walls have 57cm. The moderate occlusion space is divided into two 27.5x78cm partitions, while the highest occlusion condition has eight 23x17.5cm partitions.

4.5.2 Apparatus

This experiment was conducted with a couple of Apple iPad Air 2 that only differ from the iPad mini 4 of the previous experiment on a screen size that is 1.8 inches larger (9.7 inches) and is 138g heavier (437g). The device replacement was necessary due to the availability of two identical devices.

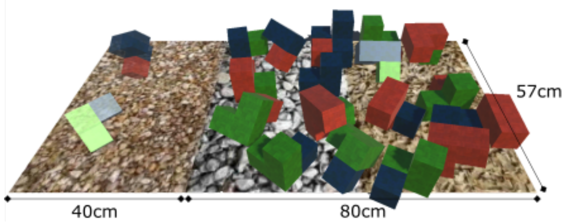
4.5.3 Subjects

Thirty subjects voluntarily took part in this experiment (thirteen female), aged 22.2 years in average (SD=2.93). They were all students with no movement restrictions on wrists and arms. We arranged the participants in pairs. They were allowed to choose their partners. Two pairs had never met before. Five subjects participated in the previous experiment (single user), two of them together.

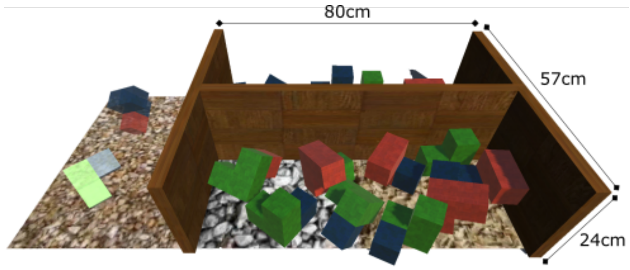
4.5.4 Experimental Setup

The experiment followed a repeated measures within-subject design with the *Occlusion (No Occlusion, Moderate Occlusion and High Occlusion)* as the independent variable.

Figure 4.12: The three occlusion conditions presented in the experiments. (a) no occlusion, (b) moderate occlusion and (c) high occlusion. On the left of each condition is the *docking space* (40cm) and on the right is the *supplies space* (80cm).



(a)



(b)

(c)

Source: the Author

The dependent variables collected were total piece manipulation *time* needed to complete each docking, and *manipulation time* of each user in each docking. Based on the results achieved in the *Experiment 1* (Sec. 4.3), we adopted the *Hybrid* manipulation during this experiment.

The participants answered a characterization form before arriving at the experiment. We explained the interface operation and allowed the users to practice the transformations and train in two docking trials. Then, the participants performed 3 blocks of 8 recorded trials. Latin squares determined the presentation order of the trial blocks with 6 different group orders. We asked the pairs to complete each docking as fast as possible. No reference errors were displayed during the recorded trials. The threshold error for a successful docking was 1.15cm in position, 8° in the rotation and 1cm in size. After each block of trials, the users answered a SEQ (see Appx. D) to assess the task difficulty and questions about behavioral interaction, mutual assistance and dependent action which encompasses the behavioral engagement factor of the *Networked Minds Measure of Social Presence* (BIOCCA; HARMS; GREGG, 2001) (see Appx. G). In the end, users filled a final form with custom questions and with questions about psychological involvement factor ((BIOCCA; HARMS; GREGG, 2001)). The experiment took on average 50 minutes.

In summary, the experiment design had: **30** participants $\times$ **3** levels of occlusion $\times$ **8** trials $=$ **720** unique dockings.

4.6 Results

4.6.1 Strategies

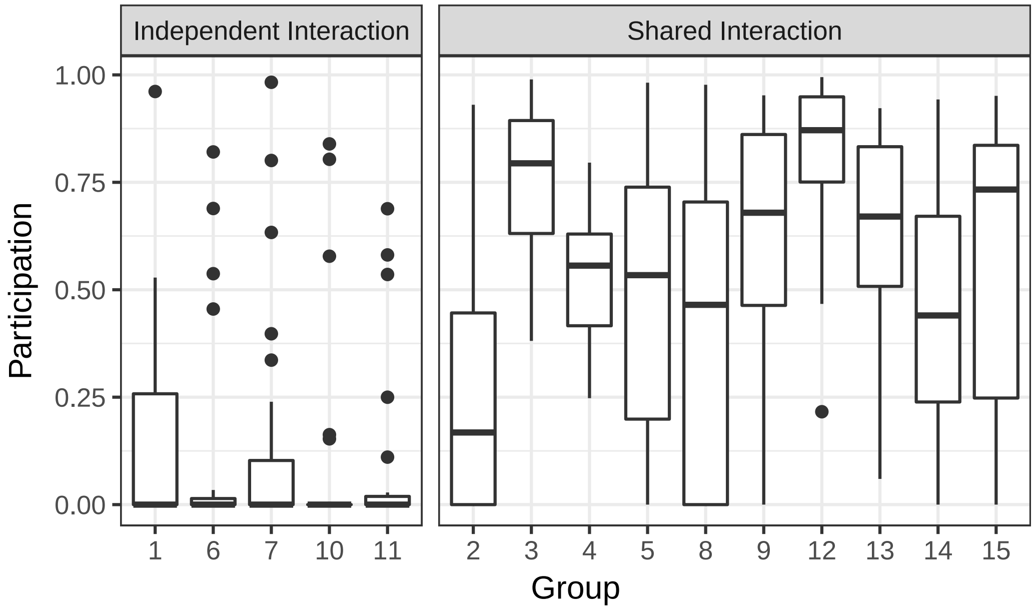
We asked participants two questions in the post-test questionnaire about the group strategy: "What was the strategy adopted by your team?" and "Has the strategy changed along the experiment?". Ten groups of fifteen (66.6%) answered the first question by saying that they choose to work together to solve the same piece, while five groups – IDs 1, 6, 7, 10 and 11 – (33.3%) said they adopted division strategy, where each piece was solved by one user alone.

We analyzed how pieces were manipulated during the trials to verify if the strategy pointed in the questionnaire was consistent with the behavior of the pair during the trials. We calculated participation score for each trial that represents the balanced participation of both subjects in the manipulation of a single piece:

$$participation = 1 - abs(\frac{Time_{User1}}{TotalTime} - \frac{Time_{User2}}{TotalTime}) \quad (4.2)$$

A score of 1 represents an equal time of manipulation, while a score of 0 means that a single user carried the task for a given piece. Figure 4.13 shows the distribution of participation scores of the pieces for each group. The Figure shows that our participation score effectively captured the work strategy reported by users.

Figure 4.13: Participation score to complete the trials for each group. Higher scores represent more simultaneous manipulations. Groups are classified by their strategy reported in the post-test questionnaire.

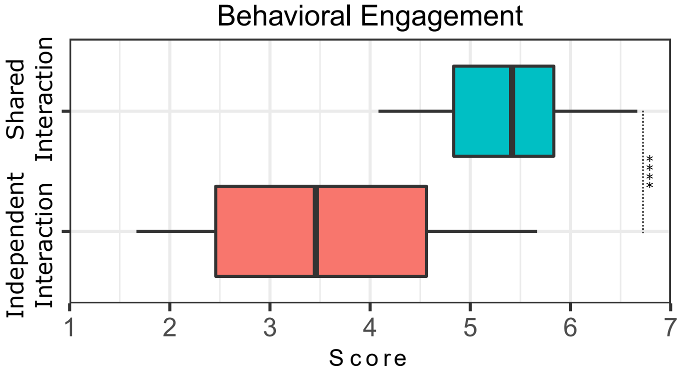


Source: the Author

Moreover, we evaluated the behavioral engagement dimension of the *Networked Minds Measure of Social Presence* (BIOCCA; HARMS; GREGG, 2001) against the two strategies adopted to observe if participants feel more engaged when docking simultaneously the same piece. Figure 4.14 shows the level of behavioral engagement of each strategy. The Wilcoxson signed-ranks test indicates a significant effect of the behavioral engagement on the strategies ($Z = 225$, $p < .001$). The ANOVA test of the behavioral engagement and participation scores shows that they are closely related ($F_{(1,13)} = 19.81$, $p < .001$), and validate our participation score as a proxy to the pair engagement while performing the experiment.

Finally, we have investigated whether the participation score explains the variation of the docking time ($TotalTime$) and the total manipulation time ($Time_{User1} + Time_{User2}$)

Figure 4.14: The users behavioral engagement in the two strategies. The results are reported in the Likert scale where 1 is low engagement and 7 is high engagement. Shared Interaction (M = 5.32, SD = 0.58), Independet Interaction (M = 3.57, SD = 1.12).



Source: the Author

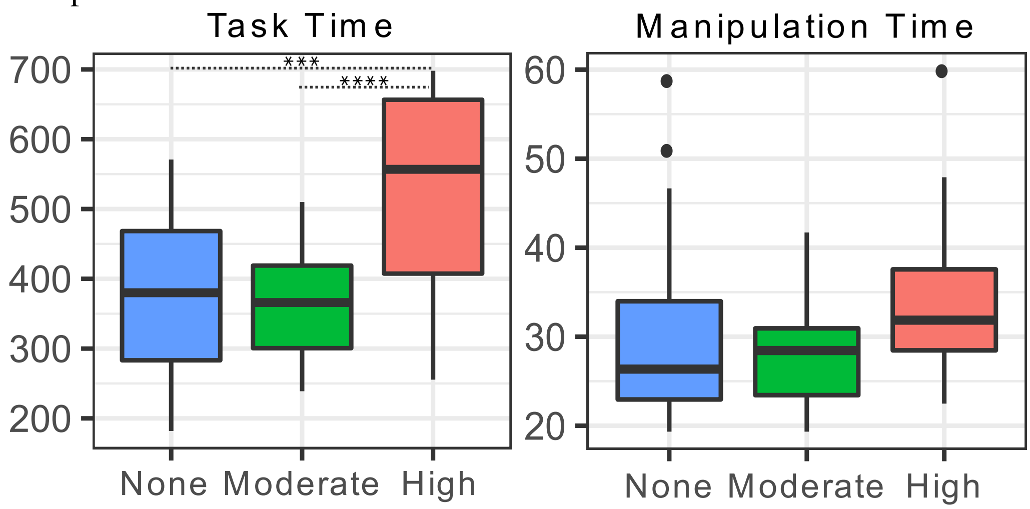
of the pieces using one-way ANOVA.The test failed to reject the equality of mean docking time and piece manipulation time across the range of computed participation scores ($F_{(1,13)} = .18$, $p > .67$ and $F_{(1,13)} = .65$, $p > .43$ respectively). That is, there is no strong evidence that the manipulation of the same piece by both users interferes with performance.

4.6.2 Occlusion vs. Task Time and Manipulation Time

We verified the relation between the independent (occlusion) and dependent (time) variables. Figure 4.15 shows the task time and manipulation trial time achieved by each occlusion condition. The ANOVA test revealed a statistically significant effect of occlusion on task completion time ($F_{(2,28)} = 9.53$, $p < .001$). The post-hoc t-test indicates significant time increase occurs for *High Occlusion* as compared to *No Occlusion* ($t_{14} = 2.9$ $p < .03$) and *Moderate Occlusion* ($t_{14} = 4$ $p < .004$). Equivalence could not be rejected for between *No Occlusion* and *Moderate Occlusion* ($t_{14} = .6$ $p > .5$).

The statistical test failed to reject equivalence in the time required to manipulate and dock the object across the levels of occlusion ($F_{(2,28)} = 1.0$, $p > .37$). This indicates that the difference in task time is due to the added search time caused by high level of occlusion.

Figure 4.15: Time vs Occlusion conditions. Task time is the time taken to complete the task. Manipulation time is the time spent by the users with manipulation commands to dock the pieces.



(a) (b)

Source: the Author

4.7 Discussion

4.7.1 Interaction Technique

The finding in our first experiment indicates that the *Hybrid* method is the most suitable for 6-DOF manipulations. We went further and observed that the time performance of the *Hybrid* is as good as the *Movements* for positioning and as good as *Touch Gestures* for rotations. This suggests that users can effectively coordinate the use of the most suitable method for each transformation and that the change between modes is made intuitively and seamlessly without the need of context-aware techniques (MOSSEL; VENDITTI; KAUFMANN, 2013). We also find that users reach similar precision with all methods when they have to indicate when they were satisfied with the docking. The precision similarity was possible because we let the users change their point-of-view. In situations where users have limited movements and need to manipulate distant objects, the *Movements* method alone is unusable (SAMINI; PALMERIUS, 2016). The lower workload reported for *Hybrid* and the SEQ corroborate with the performance analysis. The highest workload was achieved with *Movements*. The effort and frustration factors significantly affected the *Movements* workload as users have to move the device to transform the objects position and orientation. Users found it easier to manipulate with *Hybrid* and harder with *Movements*, while *Touch Gestures* received an intermediate score. *Touch Gestures* was the slowest to reach the threshold.

4.7.2 Pair Work Strategies

As hypothesized, we have observed that pairs adopted two main strategies. The 66.6% of the pairs manipulated the objects simultaneously, which we called *Shared Interaction*. The remaining pairs manipulated the objects individually, which we called *Independent Interaction*. The time performance between the two groups was similar regardless of the strategy. The experience of the participants with non-conventional devices was also similar between groups: M=2.35 SD=1.1 for the *Shared Interaction* group and M=2.35, SD=0.8 for the *Independent Interaction* group. It suggests that the user experience with non-conventional devices did not affect the strategic decision. Moreover, groups did not change their behavior with the increase of occlusion. These results indicate that the strategies are less related to the environmental factors and more related to the users and pairs profile.

We investigated the participants' behavioral engagement in the different strategies. The behavioral engagement is the degree users believe that their actions are interdependent, connected or in response to the other's actions (BIOCCA; HARMS; GREGG, 2001). The results showed that pairs that adopted the shared interaction felt more involved in the task than pairs that adopted the independent strategy, even though, in both cases, they were working together.

A collaborative interface provides each user with an individual action perspective of the same scene. The individual viewpoints allow users to place themselves on key locations, avoiding the need for constant movement to check occluded parts. Groups may adopt different strategies to complete the tasks depending on their profile and task requirements, while maintaining similar precision and time performances.

4.8 Chapter Summary

In this chapter, we presented the design of a novel user interface for collaborative manipulation of 3D augmentations superimposed on the physical environment. The technique was designed for handheld devices as they are versatile and ubiquitous in the everyday tasks. We designed two modes for 3D manipulations, touch gestures and device movements, which combined allow for intuitive 7-DOF transformations. Thanks to the *Sum-of-Contributions* model, the technique creates a shared medium where multiple users can simultaneously interact with their own devices. The technique solves the problem of concurrent manipulations with an action coordination approach, where every contribution from each user counts as a transformation step and is applied directly to the object.

More than fifty participants have tested the interface in public demonstrations. They could download the app in their smartphones and join on-going demos. Up to five team-members using devices of different models were registered during the demos. All of them easily and quickly understood the purpose and mechanics of the technique, indicating a high affordance. The observation of groups during the demos inspired the design of the two experiments.

We demonstrated the effectiveness of the *Hybrid* approach when compared with solely *Touch Gestures* or *Movements* methods in the interface assessment. Moreover, we observed that users could seamlessly switch between methods and use the most efficient action to correctly transform the object while keeping high time performance. In the second experiment with pairs, we observed that two strategies were adopted when we do not impose any restriction to collaboration. Interestingly, pairs with different strategies obtained similar task performance while teams that privileged *Shared Interaction* strategy felt more engaged in the task.

We reported the design, assessment and results presented in this Chapter in two papers. The first paper presented the initial interface design and was published as part of the IEEE 3D User Interfaces Contest in 2017 (GRANDI et al., 2017a). Our solution won the first prize among six other contenders. We published the collaborative evaluation results in the IEEE Virtual Reality Conference in 2018 (GRANDI et al., 2018) where we won an honorable mention for best paper of the conference.

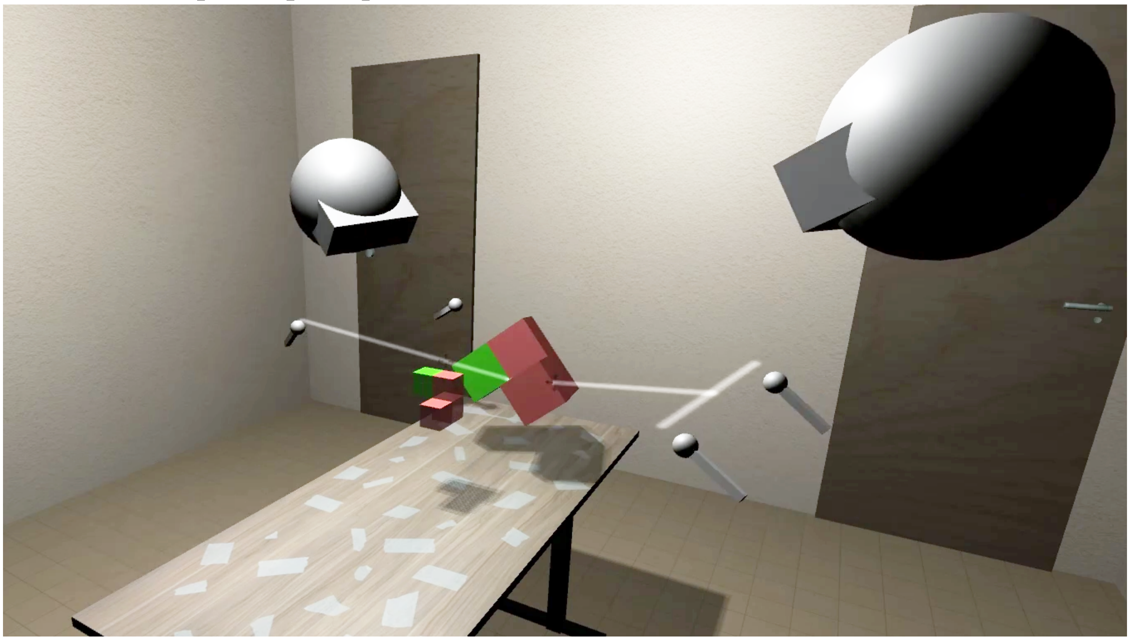
5 COLLABORATIVE 3D MANIPULATIONS FOR IMMERSIVE VIRTUAL REALITY

Interaction in immersive VR typically uses the *Virtual Hand* metaphor (JR et al., 2017) to direct map the user's hand movements to their virtual hand. While the *Virtual Hand* is the most natural way to interact with 3D objects, the interaction space is constrained by the arm's reach. Selecting and manipulating objects out of reach usually requires user re-positioning that can be by walking or through navigation techniques. Another problem it that its egocentric nature (the object is attached to the hand) makes it harder for simultaneous manipulations by multiple users. To cope with the constraints of *Virtual Hand* and allow for collaborative co-manipulations, we designed an interface to manipulate 7 DOFs on objects beyond the arms reach. The interactions may occur with one hand or combining both hands with synchronous movements. We implemented a variation of the *HOMER (Hand-centered Object Manipulation Extending Ray-Casting)* technique (BOWMAN; HODGES, 1997) for one-hand manipulations and the two-handed manipulations were based on the *Spindle* technique (MAPES; MOSHELL, 1995) for synchronous bimanual interaction. Moreover, our technique was designed using the *Sum-of-Contributions* model for multiple concurrent and synchronous modification of the same DoF by an unlimited number of users (Level 3.2 on the Margery, Arnaldi and Plouzeau (1999) classification). The technique is implemented to work on commercial HMDs controllers, such as the Oculus Rift and the HTC Vive.

5.1 3D Manipulations

A previously selected object is handled regardless of the distance that they are from the user's virtual hand. The *trigger* button is used to *grab* the object. In single hand manipulations, while the button is pressed, translation and rotation are applied directly on a 1:1 mapping. In this way, manipulation of objects in reach works similarly to what occurs in the *Virtual Hand* technique (Figure 5.2). Besides that, with our technique users can use *clutch* to move the object away or closer or to reach a total rotation beyond the wrist limits (i.e., press the trigger, move, release the trigger, hand repositioning, repeat). Additionally, the slide of the finger up and down in the joystick uniformly modify the object's scale. The scale increments can vary depending on the position of the joystick (Figure 5.2b).

Figure 5.1: Two users are simultaneously interacting with a 3D object in immersive VR. The technique allows for manipulations beyond the arm's reach. Rays drawn from the hand to the object represent the selection. Users can either choose to interact with one or two hands. Simple shapes represent the user's avatars in the VE.

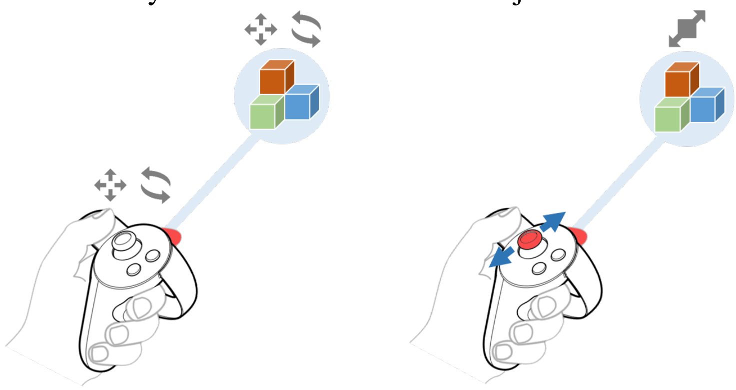


Source: the Author

Bimanual interaction is activated whenever the *trigger* button is pressed and held in both hands. In this mode, the hands perform symmetric-synchronous movements to execute the transformations (ULINSKI et al., 2009). The move of the two hands in the same direction translates the object (Figure 5.3a). Moving the hands in opposite directions (up, down, forward, backward) rotates the object in the y and z axis (Figure 5.3b). Pitch rotation is achieved by pitch rotating the hands front and back. This rotation around the x axis is the main improvement from the original *Spindle* technique (MAPES; MOSHELL, 1995) symmetric-synchronous interactions. Finally, moving the hands away or closer to each other changes the object scale (Figure 5.3c). We adopt the symmetric-synchronous interactions since it outperforms the other approaches for bimanual manipulations (ULINSKI et al., 2009). In this way, the asymmetric-synchronous variant of the *Spindle* technique proposed by Cho and Wartell (2015) was discarded.

All three transformations can be executed simultaneously in both single and bimanual interaction. It is possible to select transformations and lock the others. Figure 5.4 shows the actions to lock specific transformations. The buttons can be re-configurable depending on the controller used, in this design we are using the Oculus Touch controllers. The locks are optional and can be combined by pressing more than one button at the same time. Moreover, the *trigger* button has priority over the lock buttons, when pressed while a

Figure 5.2: Interaction gestures for distant manipulations with one hand. While pressing and holding the trigger button with the index finger it is possible to translate, rotate. The manipulations metaphor is similar to the *HOMER* technique (BOWMAN; HODGES, 1997). In addition to the classical technique, we allow the modification of the scale parameter using the joystick to uniformly increase or decrease the object's size.



(a) Translation and Rotation (b) Scale

Source: the Author

lock button is pressed, it unlocks all transformations. The transformations lock resource add more strategy possibilities during simultaneous manipulations. The groups can define roles during the manipulation task in order to separate the DOFs.

5.1.1 Simultaneous Manipulations

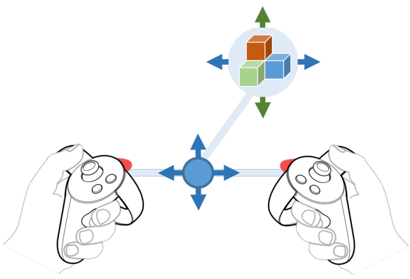
Our technique uses the *Sum-of-Contributions* (section 2.4.1) approach to handle synchronous simultaneous manipulations. The model works integrating the transformation actions performed by each user into the final object's transformation without restrictions or weights (see Chapter 2.4.1).

We added virtual avatars to make users aware of the others. Every time that the user manipulates the object, we draw virtual rays from the hand that is interacting to the selected object. Iconic representations inform which transformation the user is performing (see Figure 5.1).

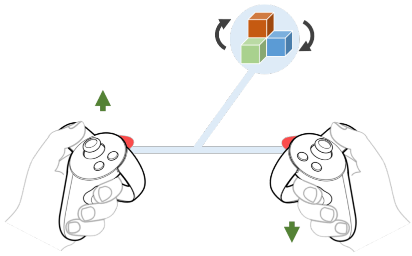
5.2 Technique Evaluation

We conducted an experiment to assess the impact on the performance of the *Distant* technique design against the benchmark *Virtual Hand* in single user manipulations.

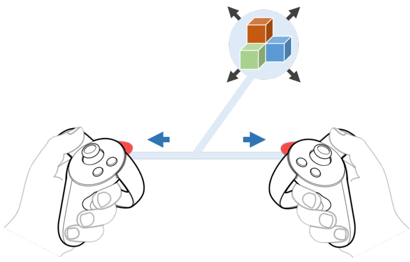
Figure 5.3: Synchronous Bimanual Interactions inspired on the work of Mapes and Moshell (1995).



(a) Bimanual Translation



(b) Bimanual Rotation



(c) Bimanual Scale

Source: the Author

Furthermore, analyzed the effect that collaborative co-manipulation have on performance and workload when compared with single user manipulations.

We investigate the following hypotheses:

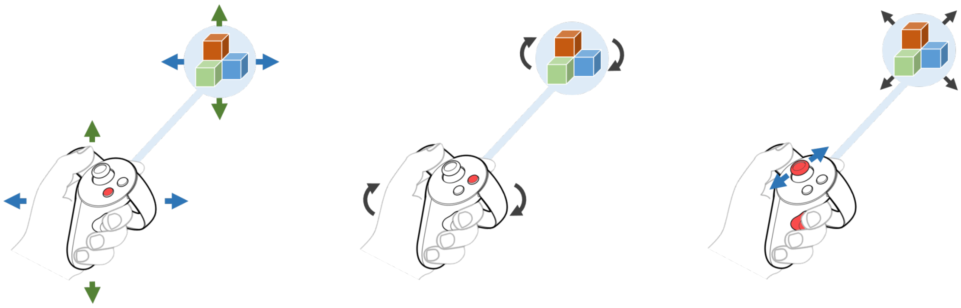
- H1. The *Virtual Hand* is faster than with the *Distant* technique in individual interaction.

- H2. Co-manipulation in pairs can achieve the same time performance of the *Virtual Hand*.

- H3. The workload is significantly lower in co-manipulation scenarios.

Even though we consider the *Virtual Hand* the most natural approach and consequently the "gold standard" for VR interaction, with this experiment we want to assess how far is our *Distant* technique from the *Virtual Hand* approach. We expect that this gap if it exists, can be filled with co-manipulations. Although previous experiments could not confirm this hypothesis. If the workload is reduced in co-manipulations, this can indicate that the work division in playing a role in the collaborative tasks.

Figure 5.4: Transformations locking allow to focus on only one transformation at a time. It is achieved by pressing specific buttons on the controller interface. It is possible to individually alter the translation (a), the rotation (b) and the scale (c).



(a) Translation Lock (b) Rotation Lock (c) Scale Lock

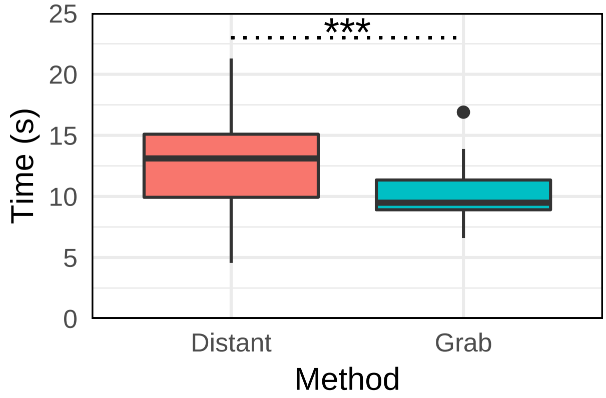
Source: the Author

5.2.1 Evaluation Design

We redesign the 3D docking task presented in Chapter 4 for our experimental setup. The docking consists of aligning the position (3 DOFs), rotation (3 DOFs) and scale (1 DOF, uniform scale) of the controlled virtual object with a similar target object. In our case, we set the target with 50% transparency to facilitate visualization when both the controlled and the target objects were interpenetrated. To recap, the cube blocks similar to the Shepard and Metzler (SHEPARD; METZLER, 1971) construction compose the objects stimuli. The blocks have 6.5cm long edges and have different colors to avoid ambiguity. The stimuli are composed by the maximum of two blocks aligned, totaling 13cm in size on each axis. The experiment was conducted using repeated measures within-subjects with the *Technique (Virtual Hand, Distant)* the independent variable for the single user evaluation. To compare the single user performance against pair work, we conducted a between-subjects with *Individual and Pairs* as the independent variable. We collected the *time* to complete each docking and users workload as dependent variables. We counterbalanced the presentation order of the conditions. After a training session, participants had to complete eight docking trials in sequence. We instructed participants to complete each trial as fast as possible. The maximum threshold tolerance for a successful docking was 1cm in position, 3° in rotation and 1% difference in scale. After each condition, while resting, the subjects answered a Single Easy Question (SEQ) (see Appx. D) to assess the task difficulty and the NASA TLX questionnaire (see Appx. F).

We analyze whether the independent variables had a significant effect in the *trial time* with a repeated measures ANOVA test. We verify whether the ANOVA assumption of

Figure 5.5: Box and whiskers plot of the trial completion time for each level of the Technique variable. Users perform faster with the Grab technique (M = 10.32, SD = 2.95) than with Distant (M = 12.69, SD = 4.32).



Source: the Author

normality of the residuals was violated with the Shapiro-Wilk test. The alpha significance level was set to 0.05. A total of twelve participants with none or very little experience with HMD VR voluntarily took part in this experiment.

5.3 Results

5.3.1 Technique Performance in Single User Interaction

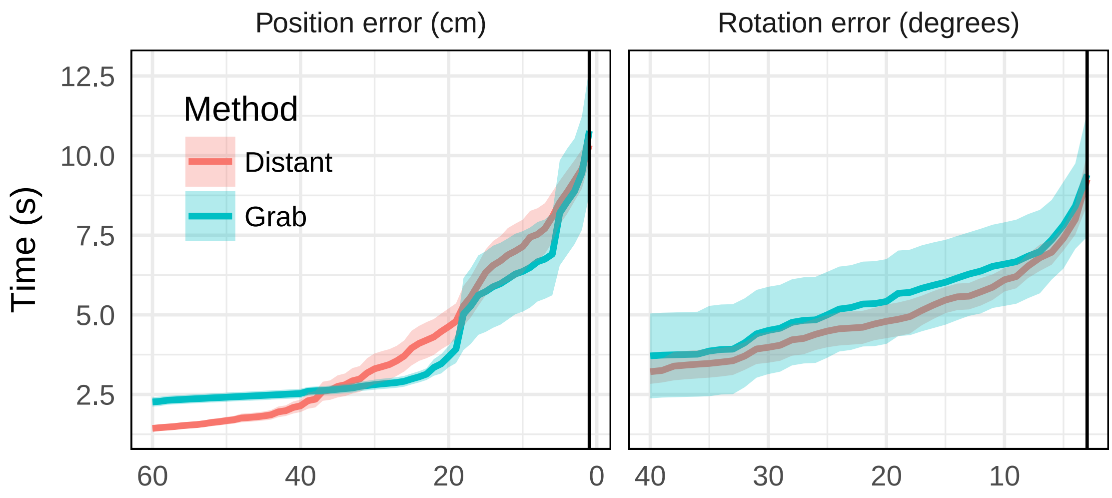
We found a significant effect for the interaction *Technique* variable ($F_{1,11} = 5.7$, $p < .04$), with *Virtual Hand* performing faster than *Distant* manipulation. Trial completion time results are presented in Figure 5.5. In Figure 5.6 we show the expected time by error reduction rate for position and rotation error of the object and its docking counterpart. *Distant* and *Virtual Hand* manipulation techniques presented similar error reduction performance.

The System Usability Score (SUS) ranged from 62.5 to 95 (M=80, SD=9.36). According to surveys that compare SUS scores for different systems, the system score is ranked as "Good" (BANGOR; KORTUM; MILLER, 2009).

5.3.2 Individual vs Collaborative Performance and Workload

Figure 5.7 shows the trial completion time for users using the Distant technique in the first experiment (individual) and the second experiment (pairs). Users interacting

Figure 5.6: Error reduction by time for position and rotation. The faster the error is reduced, the better. The shaded areas represent the standard error of the mean.



Source: the Author

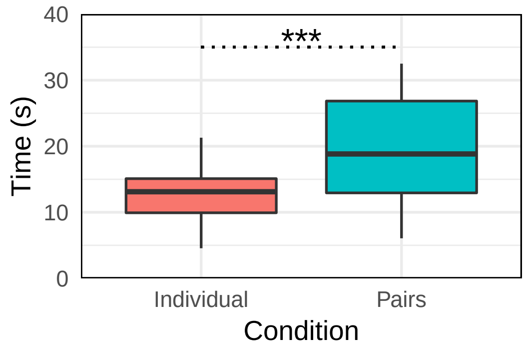
collaboratively took significantly more time to complete the trials ($p < .01$). On the other hand, the NASA TLX post-questionnaire indicates that users felt a smaller workload when interacting in pairs ($p < .02$).

5.4 Discussion

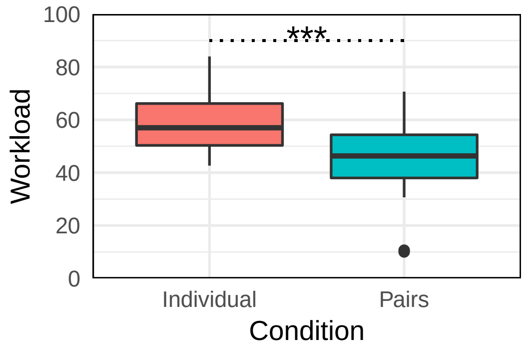
We reported the results of an experiment that compared the *Virtual Hand* with our *Distant* approach. The *Virtual Hand* is considered the most natural interaction paradigm (Bowman A. D; Kruijff, 2005), but it has limitations that restrict its use for collaborative activities. As hypothesized, it is faster to dock virtual objects with the *Virtual Hand* approach. On the other hand, the *Distant* approach had a lower variance in error during the trials, mostly in rotations. This indicates that bimanual manipulations, only available in this condition, can adjust rotations with better precision. Users reported that with more training the *Distant* would outperform the *Virtual Hand* both in performance and usability for everyday tasks similar to the Schultheis et al. (2012) conclusion. The SUS score reported reinforces this statement.

The results showed that working in pairs did not lead to an increase in speed during the task resolution. This behavior corroborates with the previous results in Chapter 5. After, we investigated whether the participant's workload is affected by working in pairs. As hypothesized, the results indicate that working in pairs can significantly reduce the

Figure 5.7: Box and whiskers plot of the trial completion time and workload for users using the Distant technique in the first experiment (individual) and in the second experiment (collaborative). (a) Individuals completed the tasks faster (M = 12.69, SD = 4.32) than Pairs (M = 19.38, SD = 8.19), (b) Individuals reported higher workload (Mdn = 57, SD = 12.76) than Pairs (Mdn = 46.33, SD = 14.76).



(a)



(b)

Source: the Author.

workload. One evidence for this behavior is the result of the participation score reported in the experiment performed in Chapter 6. The experiment reveals that both team members constantly participate in solving the manipulation task.

5.5 Chapter Summary

In this Chapter, we extended the concepts of collaborative interactions for immersive VR. First, we design a user interface capable of dealing with simultaneous manipulations in the same virtual object. The technique was designed for HMDs headsets and their controllers as they are the most intuitive and immersive hardware available at present. We proved the versatility of our *Sum-of-Contributions* model in the development of a *Distant* interaction technique. Our approach overcomes the *Virtual Hand* egocentric limitation in exchange of overall lower performance in time. The advantages of our design become evident in collaborative work. Even though the performance in pairs was worse than individuals, the significantly lower workload indicate that users collaborate in a way that the amount of work is distributed among the participants, similarly to the finds of the group sizes study present in Chapter 3.

6 ASYMMETRIC COLLABORATION WITH MIXED 3DUI TECHNIQUES

Figure 6.1: Asymmetric simultaneous interaction with mixed 3DUI techniques. Two co-located users are simultaneously interacting with a 3D object using different interfaces while sharing the same physical space.

Source: the Author

In this chapter, we present the assessment of asymmetric interactions in Collaborative Virtual Environments (CVEs). In our asymmetric setup, two co-located users interact with virtual 3D objects in immersive Virtual Reality (VR) and mobile Augmented Reality (AR), as shown in Figure 6.1. We recreated a virtual room based on the real environment to enhance Mixed Reality (MR) collaboration and the sense of co-presence. Thus, both VR and AR coexist in the same environment while keeping their view perspective, scale and dimension. While individual interactions in asymmetric MR environments are already explored to some extent (see Section 2.2), here we investigate the effectiveness of simultaneous manipulation when both users are manipulating the same 3D object. For that, we conducted an experiment to explore collaborative asymmetric interactions in a shared VR and AR environment when two users are co-manipulating virtual objects. Moreover, we want to measure the impact asymmetric interactions have on performance, social interaction and awareness. For that, we compared two *Symmetric* setups, where the interaction and visualization interfaces are the same for both participants, with the *Asymmetric* setup, where interaction and visualization are distinct for each participant. We choose the collaborative AR interface proposed in Chapter 4 for the *AR-AR Symmetric* setup. In the *VR-VR Symmetric* setup both participants use the VR technique presented in Chapter 5 and in the *VR-AR Asymmetric* setup, one participant uses the AR technique while the other uses the VR technique.

To allow VR and AR users to coexist in the same shared VE. We recreated the physical room in the virtual world. One advantage of such a setup is that co-located users can communicate with accurate positional sound. A manual calibration procedure is performed to align the physical world with the VR and AR environment. The calibration is only necessary for the VR setup since AR works with a reference marker. It consists in positioning the participant in a specific position in the physical space and move the VE. After the alignment the participant can touch and feel all real objects, such as walls and a table (passive haptics). We applied the same awareness elements present on the previous implementations, such as, the selection rays and manipulation icons. (see Chapters 3 4). These visual feedbacks are synchronized in a way that every participant knows each other's selections and actions.

Regarding performance, we hypothesize that the *VR-VR* to be the fastest, due the more direct mapping of the actions that VR provides. The *AR-AR* to be the slowest and the *VR-AR* to be in between the two other conditions. While in social interaction, we expect that in the asymmetric condition, even if participants experience the VE in different realities, they can be aware of each other, communicate and create strategies that take advantage of each interface.

6.1 Assessment of Collaborative Asymmetric Interactions

6.1.1 Task and Stimuli

The task and stimuli used in this evaluation are similar to the setup presented on the experiment in Chapter 5. All parameters, trials were kept the same and pairs manipulate the objects at the same time.

6.1.2 Aparatus

The apparatus for visualization and interaction is composed of a pair of VR headsets and a pair of AR handheld devices. The VR headset used is the *Oculus Rift* with *Touch Controllers*. Each headset had its own tracking sensors and was connected to its own

PC. Both machines had a similar configuration. The tracking sensors were positioned to avoid occlusions by the other subject. The AR device is an Apple iPad Air 2 with 9.7 inches screen ($\approx$ 264ppi density) and weights 437g. The Vuforia SDK performs the AR tracking. We track the mobile device physical pose relative to a fiducial marker. We used the table's surface as the reference marker and the extended tracking feature to extend the interaction range. Besides the interaction hardware, we used a dedicated server to manage the experimental parameters and to record all user interactions during the experiment, a dedicated network with a wired and WiFi connection. Both server and client's application were developed using the Unity3D game engine, and the communication is made with the Unity UNET network API.

6.1.3 Subjects

Thirty-six subjects participated voluntarily (11 female), aged 21.6 years on average (SD=2.19). We asked participants to choose their partner beforehand since the experiment was conducted in pairs. All subjects read and agreed to an Informed Consent Form before the experiment (see Appx. C). The majority had none or minimal experience with VR headsets, and only two subjects reported high experience. On the other hand, five subjects reported low experience with handheld tablets. They all had either normal or corrected to normal vision. Subjects with corrected vision were instructed to wear their glasses during the VR condition.

6.1.4 Experimental Setup

The experiment was conducted using repeated measures within-subjects with the *Technique Symmetry (VR-VR, AR-AR and VR-AR)* and *Co-manipulation (weather manipulating the same or different objects)* as independent variables. We collected the *time* to complete each trial, and the *manipulation time* of each subject as dependent variable. The conditions are illustrated in Figure 6.2.

Latin squares determined the presentation order of the three conditions. It resulted in 6 different orders. We explained the operation of the interface at the beginning of each condition. Subjects practiced the manipulations and performed three training trials. Then, they had to complete eight docking trials in sequence. We asked the pairs to complete each

Figure 6.2: Symmetric and Asymmetric conditions. All conditions shared the same environment configurations.

(a) Symmetric AR (b) Symmetric VR (c) Asymmetric VR and AR

Source: the Author

trial as fast as possible but avoiding to rush, that could lead to inaccuracy. The maximum threshold tolerance for a successful docking was 1cm in position, 3° in rotation and 1% difference in scale. After each condition, while resting, the subjects answered a single easy question (SEQ) to assess the task difficulty and questions of the *Networked Minds Measure of Social Presence* (BIOCCA; HARMS; GREGG, 2001) questionnaire about behavioral interaction, mutual assistance and dependent action. In the end, users filled a final form with questions about the psychological involvement factor ((BIOCCA; HARMS; GREGG, 2001)) and were asked to compare and choose the preferred condition (see Appx. G). The experiment took on average 60 minutes. In summary, the experiment design had: **18** pairs × **3** conditions × **8** trials = **432** unique trials.

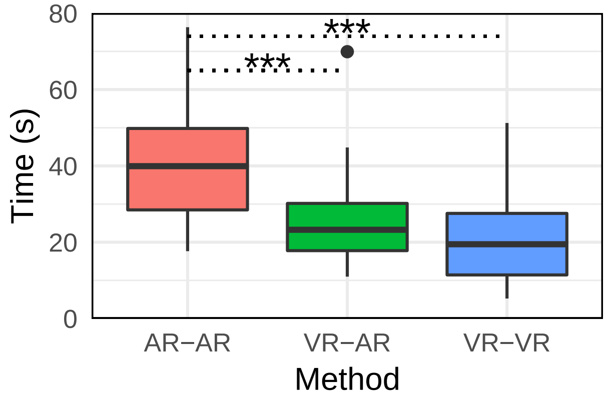
We analyze whether the independent variables – *Technique Symmetry* and *Co-manipulation* – and their interaction had a significant effect in the trial time with a 2-way repeated measures ANOVA test. We verify whether the ANOVA assumption of normality of the residuals was violated with the Shapiro-Wilk test. If a variable with more than two levels was found statistically significant, we conducted a Post-hoc analysis with multiple pairwise t-tests of the different variables. The alpha significance level was set to 0.05.

6.2 Results

6.2.1 Trial completion time

The statistical analysis indicate the the effect of Technique Symmetry was statistically significant ($F_{2,36} = 21$, $p < .001$), while it failed to reject equality of the levels

Figure 6.3: Box and whiskers plot of the trial completion time for each level of the Technique Symmetry variable. VR-VR condition is the fastest (M = 21.21, SD = 11.61) followed by VR-AR (M = 26.04, SD = 11.45) and AR-AR (M = 41.75, SD = 16.40).



Source: the Author

of Co-manipulation ($F_{1,18} = 2.93$, $p > .1$) and the interaction between the independent variables ($F_{2,36} = 1.13$, $p > .3$). The post-hoc analysis comparing the levels of Technique Symmetry suggests that the groups performed the task significantly faster in the VRVR condition as compared to the ARAR ($p < .001$) and VR-AR ($p < .05$) conditions. Besides, groups also performed faster in the VR-AR condition than in the ARAR condition ($p < .001$). Trial completion time results are presented in Fig. 6.3.

6.2.2 Error vs. Time Trade-off

Figure 6.4 shows the expected time by error reduction rate for position and rotation of the object and its docking counterpart. The smaller amount of time needed to reduce the error suggests that VR-VR is more efficient for position and rotation. We note that the VR-AR (asymmetric) condition presented similar performance to VR-VR.

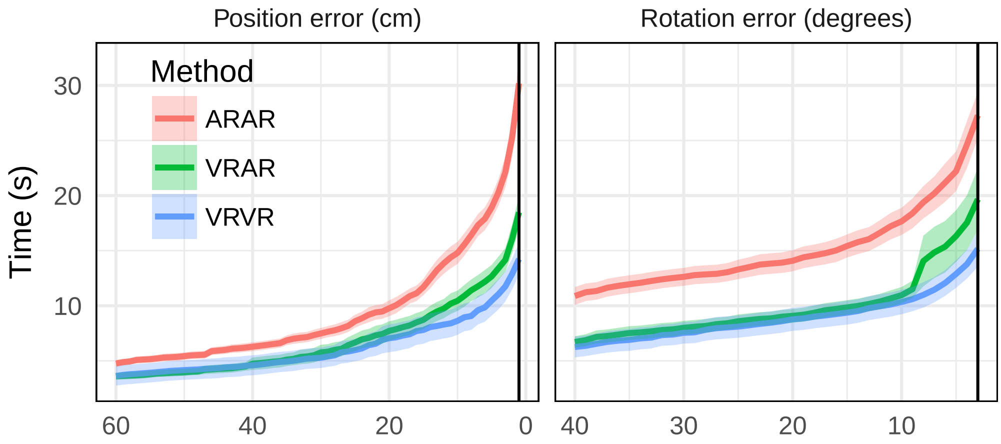
6.2.3 Participation

We estimate a participation score of group participants based the equivalence of manipulation time of each user. The participation score is computed by:

$$Participation = 1 - abs(\frac{Time_{user1} - Time_{user2}}{Time_{user1} + Time_{user2}}), \tag{6.1}$$

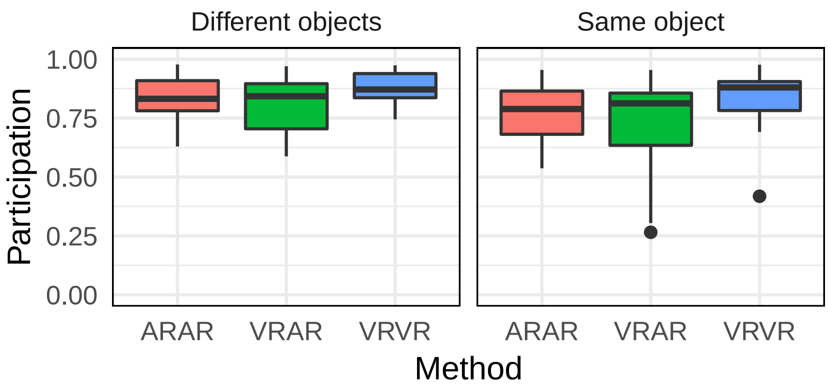
A score of 1 represents an equal time of manipulation, while a score of 0 means

Figure 6.4: Error reduction by time for position and rotation. The faster the error is reduced, the better. The shaded areas represent the standard error of the mean.



Source: the Author

Figure 6.5: Box and whiskers plot of the participation score according to the Technique Symmetry and Co-manipulation independent variables. There is no significant difference between conditions.

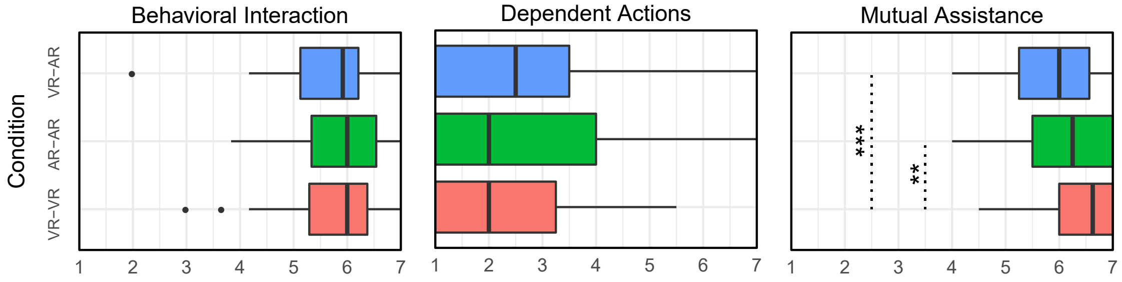


Source: the Author

that a single user carried the task for a given piece.

Statistical analysis of the Participation score shows that Co-manipulation had a significant effect ($F_{1,18} = 7.5$, $p < .02$), with improved participation time balance when users had to control different objects. Moreover, the effect of Technique Symmetry was also significant ($F_{2,36} = 4$, $p < .03$), with users in the VR-VR condition demonstrating improved participation time balance as compared to VR-AR ($p < .02$) and AR-AR ($p < .04$). Finally, no significant result was found for the interaction of the two variables ($F_{2,36} = .62$, $p > .5$). Participation results are presented in Figure 6.5.

Figure 6.6: Social presence evaluation. There is no significant difference between conditions in Behavioral interaction and Dependent actions. VR-VR has higher mutual assistance score (Mdn = 6.625, SD = 0.68), followed by AR-AR (Mdn = 6.25, SD = 0.96), and VR-AR (Mdn = 6, SD = 0.9).



Source: the Author

6.2.4 Social Presence

We evaluated the social presence with the *Networked Minds Measure of Social Presence* (BIOCCA; HARMS; GREGG, 2001). Non-parametric Friedman tests were conducted to compare the effect of the three conditions, VR-VR, AR-AR and VR-AR against three factors of the behavioral engagement dimension of Social Presence: behavioral interaction, dependent actions and mutual assistance. The post-hoc analysis was conducted with the Wilcoxon signed-ranks test with a Holm-Bonferroni correction for multiple comparisons. Figure 6.6 shows the results for each social presence dimensions. The result of the mutual assistance revealed a significant effect between conditions ($X^2(2) = 6.552, p < .037$). The post-hoc test indicates that a significant mutual assistance increase occurs from AR-AR to VR-VR ($p < .01$) and from VR-AR to VR-VR ($p < .001$). The differences occur between VR-VR and AR-AR ($p > .5$). The test failed to reject difference in independent actions ($p > .2$) factor and behavioral interaction ($p > .04$) factor.

6.3 Discussion

As hypothesized, we observed that the VR-VR outperforms both the AR-AR and VR-AR conditions. This result was expected due to the more intuitive interaction mapping that the VR controllers provide. The conditions' performance followed a consistent behavior with the asymmetric VR-AR being the second fastest and AR-AR in the third position. In this analysis, we have only evaluated the contribution and coordination behaviors. They are listed by Montoya, Massey and Lockwood (2011) as the essential collaborative behaviors that affect the team performance. We did not evaluate the communication behavior in our experiments. We investigated the user's contribution during the docking tasks to understand if, while in asymmetric, users tend to participate less when they are in interaction disadvantage. The observed results show that the team members participation remained similar in all conditions. The participation score shows that both users actively contribute to the task resolution. It strengthens the observed performance behavior trend. In the social presence analysis, we observed a significantly lower mutual assistance in the asymmetric condition. The result suggests that the different virtuality the users are exposed led to a lower sense that the pairs were helping each other.

6.4 Chapter Summary

In this Chapter, we proposed an asymmetric approach for co-manipulation in a shared environment. VR and AR users coexist in the same physical space while keeping their view perspective, scale and dimensionality. Along with the approach design, we carried out a user study to compare asymmetric and symmetric interactions in a collaborative VE. We vary the reality users experience to assess performance, participation and social presence. The user interfaces selected was the AR and VR proposed in Chapters 4 and 5 respectively. The experiment was conducted in pairs in two symmetric conditions, VR-VR and AR-AR and an asymmetric condition VR-AR. We demonstrated that pairs in VR-AR asymmetric achieved better performance than the AR-AR condition, and slightly worse performance than VR-VR. In all situations, participants had similar work participation both when they were manipulating the same or different objects. This indicates that pairs are cooperating to solve the task. Even though the users reported lower mutual assistance in asymmetric setup, we observed that the participants frequently talk about their actions and strategies during the experiment. That was also observed in the post-experiment questionnaires, where several users reported the adoption of a work division strategy. Work division strategies were also observed in previous analyses, Chapter 4.

7 FINAL CONSIDERATIONS, CONCLUSION AND PERSPECTIVES

In this thesis, we presented research that focused on simultaneous manipulations of 3D objects in collaborative virtual environments. We aimed at the design of a robust model capable of handling inputs from several sources, which we called *Sum-of-Contributions* (SoC). We integrated the SoC model into the design of 3D user interfaces for collaboration in mixed reality environments. We took into account the input hardware characteristics, the display capabilities, and awareness needs during the design process. We developed our solutions taking into account the best practices present in the literature. Along with the interface design, we assessed the users' behavior during collaborative tasks. Below we summarize the phases reported in this thesis:

1. We explored the use of handheld devices as a controller for a second screen. In this scenario, a group of users co-manipulated a virtual object while sharing the same screen and point-of-view. We investigated the relationship between group sizes and the time and accuracy to complete manipulation tasks;

2. In a second phase, we extended the interface for collaborative manipulations in augmented reality environments. In mobile AR, the screen is shared by the visualization and interaction, posing new challenges. As a benefit, users have their personalized point-of-view depending on their position in the environment. We evaluated three alternative interaction designs and assessed the capability of pairs to solve docking task letting them choose their collaborative strategies;

3. We applied the SoC model for immersive VR collaborative interactions. First, we compared the *Virtual Hand* approach, that is the most natural interaction for immersive VR interaction, with our *Distant Interaction* based on the *HOMER* approach in a single user evaluation. Then, we compared the performance and workload of pair work against single users;

4. In the end, we integrated the AR and VR solutions into an asymmetric collaborative environment. We compared performance and collaboration aspects of the asymmetric approach against two symmetric designs where pairs interacted either only in augmented reality or only in virtual reality.

7.1 Contributions

1. The Sum-of-Contributions, a model to support asymmetric inputs from different sources for fluent simultaneous manipulations.
2. A handheld-based interface for collaborative manipulations on a second screen;
3. A mobile interface for collaborative manipulations in augmented reality;
4. An interface for collaborative manipulations in immersive virtual reality;
5. A design of collaborative interactions for asymmetric mixed reality;
6. The findings of the influence of group sizes in collaborative tasks;
7. A user study to evaluate interface designs for 3D interactions in augmented reality;
8. A user study to evaluate interface designs for 3D interactions in virtual reality;
9. The findings of the influence of group strategies in collaborative tasks;
10. The findings of the influence of device asymmetry in collaborative tasks.

7.2 Design Considerations and Insights

Here we wrap up a set of design considerations and insights into the essential points approached in this thesis.

Collaboration can be beneficial in 3D tasks where precision is necessary. If the task requires speed over precision, it is necessary to take into consideration the interfaces involved and how users are integrated into the environment to solve the task. In our studies, we could not find any advantage in speed in collaboration over single users. Regarding the number of participants simultaneously manipulating the same object, we observed that adding more members in a group can cause a drop in individual productivity (Ringelmann effect). While further studies about group sizes are necessary, we would recommend keeping the group size with the maximum of four participants manipulating the same object. We also noticed that the participation during the tasks is balanced among participants. So, it is expected that if the pairs are well trained, they can equally contribute to solve the taks even in asymmetric setups.

In all studies, we observed that users plan strategies to solve the manipulation tasks, even if participants do not know each other. Participants did not report any discomfort in working with strangers. If simultaneous manipulations are mandatory, we found out that a work division strategy resulted in the best performance than other strategies. In situations where pairs could choose whether they manipulate the same or different objects to solve a task, we observed that both approaches have similar performance. However, pairs that worked together on the same object simultaneously felt more engaged during the tasks.

Except for tangible user interfaces (TUIs), the indirect interaction approach is best suitable for collaborative manipulations. Even though the direct approach, using the virtual hand technique, for example, is the most natural interaction, it is rather egocentric and not friendly for simultaneous manipulations. On the other hand, 3D manipulations at a distance, while less natural and slower than a direct approach, can be implemented taking advantage of simultaneous inputs from different sources at the same time while keeping good affordance.

We highlight that providing visual feedback of the selections and actions of the other are essential for user's awareness during the collaborative task, even in co-located collaboration where communication was not an issue. The capacity to show what they are doing improves the knowledge about the task and the others.

7.3 Applications

We envision the application of collaborative manipulations in several areas. In interior design, such as the work of Ibayashi et al. (2015). In an asymmetric setup, it would be possible for two designers to share the same virtual room while in different virtualities. For instance, one could be immersed in VR interacting with the virtual furniture in a simulated room, while the other could be in the real room and interact with augmentations of the furniture in real scale. The application would allow both designers to visualize and communicate with each other inside the room while using the simultaneous manipulation technique to make to the room contents. Collaborative manipulations could help professionals that work with large, complex 3D (or higher-dimensional) datasets. Several users could visualize the data in different scales and immersion, and spatially manipulate the datasets using the interfaces proposed in this thesis. The users could immerse through the data. This could lead to an understanding and insight that could not be obtained only looking at numeric results. Collaborative interfaces could also be very effective in training and teaching tasks such as assembling, prototype building, and architectural creation.

7.4 Thesis in Numbers

- 158 subjects voluntarily participated in our user studies;
- 100 hours were spent in experimental sessions;
- Our interfaces were presented in four public demonstrations;
- Over two-hundred people tested our interfaces in demos;
- Over 3GB of usage data were generated during the experiments and analyzed to report the results;
- Four papers published;
- Two prizes for best 3D User Interface;
- One honorable mention of Best Paper.

7.5 Perspectives

Our research contributes to several aspects of collaborative interactions. Even so, we left several possible open research tracks open.

First, this thesis presented studies for fundamental research questions. Thus, we evaluate all designs with classical 3D interaction tasks. Further studies would encompass the investigation of collaborative aspects during interactions in real case scenarios. Evaluations in real applications would raise new hypotheses and give new insights about users behavior during collaboration.

Second, we narrowed our study only for object manipulations in 3D virtual environments. We also provided means for selections, but we only use concepts already established. Selection is one of the canonical actions in virtual environments and can be further explored during collaborative interactions. Collaborative navigation is another research venue that our thesis can benefit. Since selections were not approached in this work, we only allow manipulations through the object's center-of-mass. Future implementations could add the possibility to select and manipulate the object through a pivot point, like the work of Duval, Lécuyer and Thomas (2006).

Third, we presented our evaluations with co-located users. Thus, the group shares the same physical space. Even though our techniques are built upon standard internet protocols, we only use them for local communications. Solutions for remote communications and evaluations about latency and other infrastructure problems are left for future investigations.

We make the *Sum-of-Contributions* the most generic model possible. Since the model sums the contributions of all sources, if two sources perform an action in the same direction, the resulting object's movement will be twice as expected. One future improvement is the implementation of an input compensation control when movements are performed in the same direction.

We limited our augmented reality setup with the use of handheld devices. Visualization hardware that works as an HMD for AR, such as the *Hololens* is a promising technology for Mixed Reality (MR). New opportunity of research, with the popularity of such MR interfaces, could include the exploration of the collaborative manipulations in MR scenarios. For that, new interaction design for precise 3D manipulations is required since the hand interactions natively provided by the *Hololens* are somewhat limited. The exploration of asymmetric setups with HMD VR and the *Hololens* AR like the recent work of Piumsomboon et al. (2018) could continue in future works.

In most of our evaluations, we assessed collaborative interaction during pair work even though our solutions are designed to handle inputs of several users at the same time. The exception was in the first experiment where we evaluate the effect of group sizes in 3D co-manipulation. In future works, the collaboration in scenarios with larger groups could be approached in more detail since it is still little explored. One of the goals of our research was to assess how groups organize themselves when they are free to make their task organization. Future research could focus on different levels of controlled labor division in a way that each user will be responsible for only one specific transformation. Forcing the role division between the peers could affect collaboration and performance (SAMINI; PALMERIUS, 2016).

Regarding asymmetric interactions, we explored the interface asymmetry dimension, where we vary the interface used in the interaction. One promising research venue is the exploration of the scale asymmetry, where the size of the users in the VE changes. The scales can be combined with the interfaces described in this thesis. The participants could arrange themselves in different organizational levels depending on the level of immersion provided by the technique's visualization screen. At the operational level, where the users are the workers, a portable display or an HMD would support the viewing of individual plans, while at the tactical level, a large display could provide for the awareness of the complete task state and formulation of a strategy that includes multiple workers. Further studies could investigate if the relationships among participants influence the group performance and the strategic decisions.

REFERENCES

AGUERRECHE, L.; DUVAL, T.; LÉCUYER, A. 3-Hand Manipulation of Virtual Objects. In: **Proceedings of the 15th Joint Virtual Reality Eurographics Conference on Virtual Environments**. Aire-la-Ville, Switzerland: Eurographics Association, 2009. (JVRC'09), p. 153–156. ISBN 978-3-905674-20-0. Available from Internet: <http://dx.doi.org/10.2312/EGVE/JVRC09/153-156>.

AGUERRECHE, L.; DUVAL, T.; LÉCUYER, A. Reconfigurable tangible devices for 3D virtual object manipulation by single or multiple users. In: **Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology**. New York, USA: ACM Press, 2010. p. 227. ISBN 9781450304412. Available from Internet: <http://doi.acm.org/10.1145/1889863.1889913http://portal.acm.org/citation.cfm?doid=1889863.1889913>.

BAI, H. et al. Markerless 3D gesture-based interaction for handheld Augmented Reality interfaces. **2013 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2013**, n. October 2013, p. 0–5, 2013.

BAI, H. et al. 3D gesture interaction for handheld augmented reality. **SIGGRAPH Asia 2014 Mobile Graphics and Interactive Applications on - SA '14**, p. 1–6, 2014. Available from Internet: <http://dl.acm.org/citation.cfm?doid=2669062.2669073>.

BALLAGAS, R. et al. The Smart Phone : A Ubiquituous Input Device. **IEEE Pervasive Computing**, v. 5, p. 70–77, 2006. ISSN 1536-1268.

BANGOR, A.; KORTUM, P.; MILLER, J. Determining what individual sus scores mean: Adding an adjective rating scale. **J. Usability Studies**, Usability Professionals' Association, Bloomingdale, IL, v. 4, n. 3, p. 114–123, may 2009. ISSN 1931-3357. Available from Internet: <http://dl.acm.org/citation.cfm?id=2835587.2835589>.

BENKO, H.; FEINER, S. Balloon selection: A multi-finger technique for accurate low-fatigue 3D selection. **IEEE Symposium on 3D User Interfaces 2007 - Proceedings, 3DUI 2007**, p. 79–86, 2007.

BERGÉ, L.-p.; DUBOIS, E.; RAYNAL, M. Design and Evaluation of an " Around the SmartPhone " Technique for 3D Manipulations on Distant Display. In: **SUI15**. [S.l.: s.n.], 2015. p. 69–78. ISBN 9781450337038.

BERGÉ, L.-P. et al. Exploring Smartphone - Based Interaction with Overview + Detail Interfaces on 3D Public Displays. **MobileHCI 2014**, p. 125–134, 2014. Available from Internet: <http://dl.acm.org/citation.cfm?id=2628374{&}CFID=606332997{&}CFTOKEN=49>.

BILLINGHURST, M. et al. Mixing realities in Shared Space: an augmented reality interface for collaborative computing. **2000 IEEE International Conference on Multimedia and Expo. ICME2000. Proceedings. Latest Advances in the Fast Changing World of Multimedia (Cat. No.00TH8532)**, v. 3, n. January, p. 1641–1644, 2000. ISSN 0-7803-6536-4. Available from Internet: <http://ieeexplore.ieee.org/document/871085/>.

BIOCCA, F.; HARMS, C.; GREGG, J. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. **4th Annual Int. Workshop on Presence**, p. 1–9, 2001. Available from Internet: <https://ispr.info/presence-conferences/previous-conferences/presence-2001/>.

BORING, S. et al. Touch projector: mobile interaction through video. **SIGCHI Conference on Human Factors in Computing Systems**, p. 2287–2296, 2010.

Bowman A. D; Kruijff, E. L. J. J. P. I. **3D User Interfaces Theory and Pratice**. Boston, MA: Ed Addison-wesley, 2005. ISBN 0-201-75867-9.

BOWMAN, D. A.; HODGES, L. F. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. **Proceedings of the 1997 symposium on Interactive 3D graphics - SI3D '97**, p. 35–ff., 1997. ISSN 03032647. Available from Internet: <http://portal.acm.org/citation.cfm?doid=253284.253301>.

BROOKE, J. SUS: a retrospective. **Journal of usability studies**, Usability Professionals' Association, v. 8, n. 2, p. 29–40, 2013.

BUTLER, A.; IZADI, S.; HODGES, S. SideSight: multi-"touch" interaction around small devices. **Uist**, p. 201–204, 2008. Available from Internet: <http://portal.acm.org/citation.cfm?doid=1449715.1449746>.

CHENECHAL, M. L. et al. When the giant meets the ant an asymmetric approach for collaborative and concurrent object manipulation in a multi-scale environment. In: **2016 IEEE Third VR International Workshop on Collaborative Virtual Environments (3DCVE)**. IEEE, 2016. p. 18–22. ISBN 978-1-5090-2138-3. Available from Internet: <http://ieeexplore.ieee.org/document/7563562/>.

CHO, I.; WARTELL, Z. Evaluation of a bimanual simultaneous 7DOF interaction technique in virtual environments. In: **2015 IEEE Symposium on 3D User Interfaces (3DUI)**. IEEE, 2015. p. 133–136. ISBN 978-1-4673-6886-5. Available from Internet: <http://ieeexplore.ieee.org/document/7131738/>.

CHURCHILL, E. F.; SNOWDON, D. Collaborative virtual environments: an introductory review of issues and systems. **Virtual Reality**, Springer, v. 3, n. 1, p. 3–15, 1998.

CLERGEAUD, D. et al. Towards seamless interaction between physical and virtual locations for asymmetric collaboration. **Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology - VRST '17**, p. 1–4, 2017. Available from Internet: <http://dl.acm.org/citation.cfm?doid=3139131.3139165>.

DEBARBA, H.; NEDEL, L.; MACIEL, A. LOP-cursor: Fast and precise interaction with tiled displays using one hand and levels of precision. **IEEE Symposium on 3D User Interfaces 2012, 3DUI 2012 - Proceedings**, p. 125–132, 2012.

DEBARBA, H. G. et al. Disambiguation Canvas: A Precise Selection Technique for Virtual Environments. In: **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**. [s.n.], 2013. v. 8119 LNCS, n. PART 3, p. 388–405. ISBN 9783642404764. Available from Internet: <http://link.springer.com/10.1007/978-3-642-40477-1{_}.>

DODDS, T. J. Collaborative Interaction in Virtual Environments. p. 126, 2009. Available from Internet: <http://etheses.whiterose.ac.uk/1372/1/dodds.pdf>.

DOURISH, P.; BELLOTTI, V. Awareness and coordination in shared workspaces. In: **Proceedings of the 1992 ACM Conference on Computer-supported Cooperative Work**. New York, NY, USA: ACM, 1992. (CSCW '92), p. 107–114. ISBN 0-89791-542-9. Available from Internet: <http://doi.acm.org/10.1145/143457.143468>.

DUVAL, T.; LÉCUYER, A.; THOMAS, S. SkeweR: A 3D interaction technique for 2-user collaborative manipulation of objects in virtual environments. In: **3DUI 2006: IEEE Symposium on 3D User Interfaces 2006 - Proceedings**. [S.l.: s.n.], 2006. v. 2006, p. 69–72. ISBN 1424402255.

FORLINES, C. et al. Direct-touch vs. mouse input for tabletop displays. **Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07**, p. 647, 2007. ISSN 10892680. Available from Internet: <http://portal.acm.org/citation.cfm?doid=1240624.1240726>.

FORSBERG, A.; HERNDON, K.; ZELEZNIK, R. Aperture based selection for immersive virtual environments. In: **Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology**. New York, NY, USA: ACM, 1996. (UIST '96), p. 95–96. ISBN 0-89791-798-7. Available from Internet: <http://doi.acm.org/10.1145/237091.237105>.

FROEHLICH, B. et al. The globefish and the globemouse: two new six degree of freedom input devices for graphics applications. **Proceedings of the SIGCHI conference on Human Factors in computing systems**, p. 199, 2006. Available from Internet: <http://portal.acm.org/citation.cfm?id=1124802>.

GIESLER, A.; VALKOV, D.; HINRICHS, K. Void Shadows : Multi-Touch Interaction with Stereoscopic Objects on the Tabletop. **SUI '14 Proceedings of the 2nd ACM symposium on Spatial user interaction**, p. 104–112, 2014. ISSN 9781450328203.

GOLDSCHWENDT, T. et al. The collaborative design platform protocol. In: **Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology - VRST '14**. New York, New York, USA: ACM Press, 2014. p. 239–240. ISBN 9781450332538. Available from Internet: <http://dl.acm.org/citation.cfm?doid=2671015.2671137>.

GRANDI, J. G. et al. Collaborative 3D manipulation using mobile phones. In: **2016 IEEE Symposium on 3D User Interfaces (3DUI)**. IEEE, 2016. p. 279–280. ISBN 978-1-5090-0842-1. Available from Internet: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7460079>.

GRANDI, J. G. et al. Collaborative manipulation of 3D virtual objects in augmented reality scenarios using mobile devices. In: **2017 IEEE Symposium on 3D User Interfaces (3DUI)**. IEEE, 2017. p. 264–265. ISBN 978-1-5090-6716-9. Available from Internet: <http://ieeexplore.ieee.org/document/7893373/>.

GRANDI, J. G. et al. Design and Evaluation of a Handheld-based 3D User Interface for Collaborative Object Manipulation. In: **Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17**. New York, New York, USA:

ACM Press, 2017. p. 5881–5891. ISBN 9781450346559. Available from Internet: <http://dl.acm.org/citation.cfm?doid=3025453.3025935>.

GRANDI, J. G. et al. Design and Assessment of a Collaborative 3D Interaction Technique for Handheld Augmented Reality. In: **2018 IEEE Virtual Reality (VR)**. [S.l.]: IEEE, 2018.

GUGENHEIMER, J. et al. ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users. **Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17**, p. 4021–4033, 2017. Available from Internet: <http://dl.acm.org/citation.cfm?doid=3025453.3025683>.

HANCOCK, M.; CARPENDALE, S.; COCKBURN, A. Shallow-depth 3d interaction. In: **Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07**. New York, New York, USA: ACM Press, 2007. p. 1147. ISBN 9781595935939. Available from Internet: <http://portal.acm.org/citation.cfm?id=1240798http://portal.acm.org/citation.cfm?doid=1240624.1240798>.

HANCOCK, M.; Ten Cate, T.; CARPENDALE, S. Sticky Tools: Full 6DOF Force-Based Interaction for Multi-Touch Tables. **Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces**, v. 11, n. 6, p. 133–140, 2009. ISSN 1467-8659. Available from Internet: <http://scholar.google.com/scholar?hl=en{&}btnG=Search{&}q=intitle:Sticky+Tools:+Full+6DOF+Force-Based+Interaction+for+Multi-Touch+Ta>.

HANCOCK, M. S. et al. Rotation and translation mechanisms for tabletop interaction. **Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems, TABLETOP'06**, v. 2006, n. January, p. 79–86, 2006.

HART, S. G. Nasa-task load index (nasa-tlx); 20 years later. In: SAGE PUBLICATIONS SAGE CA: LOS ANGELES, CA. **Proceedings of the human factors and ergonomics society annual meeting**. [S.l.], 2006. v. 50, n. 9, p. 904–908.

HENRYSSON, A.; BILLINGHURST, M.; OLLILA, M. Virtual object manipulation using a mobile phone. In: **Proceedings of the 2005 international conference on Augmented tele-existence - ICAT '05**. New York, New York, USA: ACM Press, 2005. p. 164. ISBN 0473106574. Available from Internet: <http://portal.acm.org/citation.cfm?doid=1152399.1152430>.

HINCKLEY, K. et al. Usability Analysis of 3D Rotation Techniques. **Proceedings of the Symposium on User Interface Software and Technology**, p. 1–10, 1997. Available from Internet: <http://dl.acm.org/citation.cfm?id=263407.263408>.

HÜRST, W.; Van Wezel, C. Gesture-based interaction via finger tracking for mobile augmented reality. **Multimedia Tools and Applications**, v. 62, n. 1, p. 233–258, 2013. ISSN 13807501.

IBAYASHI, H. et al. Dollhouse VR: A multi-view, multi-user collaborative design workspace with VR technology. **SIGGRAPH Asia 2015 Posters, SA 2015**, p. 2–3, 2015. ISSN 1882-7764.

INGHAM, A. G. et al. The Ringelmann effect: Studies of group size and group performance. **Journal of Experimental Social Psychology**, v. 10, n. 4, p. 371–384, 1974. ISSN 10960465.

JR, J. J. L. et al. **3D user interfaces: theory and practice**. [S.l.]: Addison-Wesley Professional, 2017.

KATZAKIS, N. Plane-Casting: 3D Cursor Control With A SmartPhone. **The 3rd Dimension of CHI, Workshop at ACM SIGCHI**, p. 199–200, 2012.

KATZAKIS, N.; HORI, M. Mobile phones as 3-DOF controllers: A comparative study. **8th IEEE International Symposium on Dependable, Autonomic and Secure Computing, DASC 2009**, p. 345–349, 2009.

KATZAKIS, N. et al. INSPECT: extending plane-casting for 6-DOF control. **Human-centric Computing and Information Sciences**, Springer Berlin Heidelberg, v. 5, n. 1, p. 22, dec 2015. ISSN 2192-1962. Available from Internet: <http: //www.hcis-journal.com/content/5/1/22>.

KHALID, S.; ULLAH, S.; ALAM, A. Optimal Latency in Collaborative Virtual Environment to Increase User Performance: A Survey. **International Journal of Computer Applications**, v. 142, n. 3, p. 35–47, may 2016. ISSN 09758887. Available from Internet: <http://www.ijcaonline.org/archives/volume142/number3/ khalid-2016-ijca-909723.pdf>.

KIM, M.; LEE, J. Y. Touch and hand gesture-based interactions for directly manipulating 3D virtual objects in mobile augmented reality. **Multimedia Tools and Applications**, Multimedia Tools and Applications, p. 1–22, 2016. ISSN 15737721. Available from Internet: <http://dx.doi.org/10.1007/s11042-016-3355-9>.

KNOEDEL, S.; HACHET, M. Multi-touch RST in 2D and 3D spaces: Studying the impact of directness on user performance. **3DUI 2011 - IEEE Symposium on 3D User Interfaces 2011, Proceedings**, p. 75–78, 2011.

KRATZ, S.; ROHS, M. HoverFlow: Expanding the Design Space of Around-device Interaction. **Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services**, p. 4:1—-4:8, 2009. Available from Internet: <http://doi.acm.org/10.1145/1613858.1613864{\%}5Cnhttp: //www.scopus.com/inward/record.url?eid=2-s2.0-70450263487{&}partnerID=tZ>.

KRATZ, S. et al. PalmSpace: Continuous around-device gestures vs. multitouch for 3D rotation tasks on mobile devices. **Proceedings of the International Working Conference on Advanced Visual Interfaces**, p. 181–188, 2012. Available from Internet: <http://dl.acm.org/citation.cfm?id=2254556.2254590{\%}5Cnhttp: //dl.acm.org/citation.cfm?doid=2254556.2254>.

LAUGHLIN, P. R. **Group Problem Solving**. [S.l.]: Princeton University Press, 2011. ISBN 9780691147918.

LENGYEL, J. et al. Real-time robot motion planning using rasterizing computer graphics hardware. **ACM SIGGRAPH Computer Graphics**, v. 24, n. 4, p. 327–335, 1990. ISSN 00978930. Available from Internet: <http://portal.acm.org/citation.cfm?doid=97880. 97915>.

LIANG, H. N. et al. An investigation of suitable interactions for 3D manipulation of distant objects through a mobile device. **International Journal of Innovative Computing, Information and Control**, v. 9, n. 12, p. 4737–4752, 2013. ISSN 13494198.

LIU, C. et al. Shared Interaction on a Wall-Sized Display in a Data Manipulation Task. **Proceedings of the 34th international conference on Human factors in computing systems**, p. 1–12, 2016. Available from Internet: <https://hal.inria.fr/hal-01275535/>.

LOPEZ, D. et al. Towards An Understanding of Mobile Touch Navigation in a Stereoscopic Viewing Environment for 3D Data Exploration. **IEEE Transactions on Visualization and Computer Graphics**, v. 22, n. 5, p. 1616–1629, may 2016. ISSN 1077-2626. Available from Internet: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7118242>.

MAPES, D. P.; MOSHELL, J. M. A two-handed interface for object manipulation in virtual environments. **Presence: Teleoperators & Virtual Environments**, MIT Press, v. 4, n. 4, p. 403–416, 1995.

MARGERY, D.; ARNALDI, B.; PLOUZEAU, N. A general framework for cooperative manipulation in virtual environments. **Virtual Environments**, v. 99, p. 169–178, 1999. Available from Internet: <http://scholar.google.com/scholar?hl=en{&}btnG=Search{&}q=intitle:A+General+Framework+for+Cooperative+Manipulation+in+Virtual+Environm>.

MARTINET, A.; CASIEZ, G.; GRISONI, L. The Effect of DOF Separation in 3D Manipulation Tasks with Multi-touch Displays. **Proceedings of the 17th Symposium on Virtual Reality Software and Technology (VRST'10)**, v. 1, n. 212, p. 111–118, 2010. Available from Internet: <http://dl.acm.org/citation.cfm?id=1889863.1889888{\%}5Cnhttp://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5444709{\%}5Cnhttp://portal.acm.org/citation.cfm?id=1>.

MARTINET, A.; CASIEZ, G.; GRISONI, L. Integrality and separability of multitouch interaction techniques in 3D manipulation tasks. **IEEE Transactions on Visualization and Computer Graphics**, v. 18, n. 3, p. 369–380, 2012. ISSN 10772626.

MARZO, A.; BOSSAVIT, B.; HACHET, M. Combining multi-touch input and device movement for 3D manipulations in mobile augmented reality environments. In: **Proceedings of the Symposium on Spatial User Interaction**. New York, USA: ACM Press, 2014. p. 13–16. ISBN 9781450328203. Available from Internet: <http://dl.acm.org/citation.cfm?doid=2659766.2659775>.

MONTOYA, M. M.; MASSEY, A. P.; LOCKWOOD, N. S. 3D Collaborative Virtual Environments: Exploring the Link between Collaborative Behaviors and Team Performance. **Decision Sciences**, v. 42, n. 2, p. 451–476, 2011. ISSN 00117315.

MOSSEL, A.; VENDITTI, B.; KAUFMANN, H. 3DTouch and HOMER-S: Intuitive Manipulation Techniques for One-Handed Handheld Augmented Reality. **Proceedings of the Virtual Reality International Conference on Laval Virtual - VRIC '13**, p. 1, 2013. Available from Internet: <http://dl.acm.org/citation.cfm?doid=2466816.2466829>.

MÜLLER, J.; RÄDLE, R.; REITERER, H. Virtual Objects As Spatial Cues in Collaborative Mixed Reality Environments: How They Shape Communication Behavior and User Task Load. **Proceedings of the Conference on Human Factors in Computing Systems**, p. 1245–1249, 2016. Available from Internet: <http://doi.acm.org/10.1145/2858036.2858043>.

NGUYEN, T. T. H.; DUVAL, T. A survey of communication and awareness in collaborative virtual environments. In: **International Workshop on Collaborative Virtual Environments**. [s.n.], 2014. p. 1–8. ISBN 978-1-4799-5217-5. Available from Internet: <http://ieeexplore.ieee.org/document/7160928/>.

NGUYEN, T. T. H.; DUVAL, T.; FLEURY, C. Guiding Techniques for Collaborative Exploration in Multi-Scale Shared Virtual Environments. **GRAPP International Conference on Computer Graphics Theory and Applications**, p. 327–336, 2013. Available from Internet: <http://hal.archives-ouvertes.fr/hal-00755313>.

NORMAN, G. Likert scales, levels of measurement and the "laws" of statistics. **Advances in health sciences education**, Springer, v. 15, n. 5, p. 625–632, 2010.

OHNISHI, T. et al. Virtual interaction surface: Decoupling of interaction and view dimensions for flexible indirect 3D interaction. **IEEE Symposium on 3D User Interfaces 2012, 3DUI 2012 - Proceedings**, p. 113–116, 2012.

OKADA, K.-i. Collaboration support in the information sharing space. **IPSJ Magazine**, v. 48, n. 2, p. 123–125, 2007.

OTTO, O.; ROBERTS, D.; WOLFF, R. A review on effective closely-coupled collaboration using immersive CVE's. **Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications - VRCIA '06**, p. 145, 2006. Available from Internet: <http://dl.acm.org/citation.cfm?id=1128947{\%}5Cnhttp://portal.acm.org/citation.cfm?doid=1128923.1128>.

PIETROSZEK, K.; WALLACE, J. R.; LANK, E. Tiltcasting: 3D interaction on large displays using a mobile device. **Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15**, p. 57–62, 2015. Available from Internet: <http://dl.acm.org/citation.cfm?doid=2807442.2807471>.

PINHO, M. S.; BOWMAN, D. A.; FREITAS, C. M. D. S. Cooperative object manipulation in collaborative virtual environments. **Journal of the Brazilian Computer Society**, scielo, v. 14, p. 54–67, 2008. ISSN 0104-6500.

PIUMSOMBOON, T. et al. CoVAR: Mixed-Platform Remote Collaborative Augmented and Virtual Realities System with Shared Collaboration Cues. **Adjunct Proceedings of the 2017 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2017**, p. 218–219, 2017.

PIUMSOMBOON, T. et al. Mini-Me. In: **Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18**. New York, New York, USA: ACM Press, 2018. p. 1–13. ISBN 9781450356206. Available from Internet: <http://dl.acm.org/citation.cfm?doid=3173574.3173620>.

POUPYREV, I. et al. Egocentric Object Manipulation in Virtual Environments: Empirical Evaluation of Interaction Techniques. **Computer Graphics Forum**, v. 17, n. 3, p. 41–52, 1998. ISSN 0167-7055. Available from Internet: <http://doi.wiley.com/10.1111/1467-8659.00252>.

REISMAN, J. L.; DAVIDSON, P. L.; HAN, J. Y. A screen-space formulation for 2D and 3D direct manipulation. **Symposium on User Interface Software and Technology**, p. 9, 2009. Available from Internet: <http://portal.acm.org/citation.cfm?id=1622176.1622190>.

RIEGE, K. et al. The bent pick ray: An extended pointing technique for multi-user interaction. **3DUI 2006: IEEE Symposium on 3D User Interfaces 2006 - Proceedings**, v. 2006, p. 62–65, 2006.

ROBINSON, M. Designing for Unanticipated Use... In: **ECSCW '93**. [S.l.: s.n.], 1993. p. 187–202. ISBN 1581136935.

RUDDLE, R. A.; SAVAGE, J. C. D.; JONES, D. M. Symmetric and asymmetric action integration during cooperative object manipulation in virtual environments. **ACM Transactions on Computer-Human Interaction (TOCHI)**, ACM, v. 9, n. 4, p. 285–308, 2002.

SALZMANN, H., J.; ACOBS, J.; FROEHLICH, B. Collaborative interaction in co-located two-user scenarios. In: **Proc. of JVRC 2009 (Joint Virtual Reality Conference of EGVE - ICAT - EuroVR)**. [S.l.: s.n.], 2009. p. 85–92.

SAMINI, A.; PALMERIUS, K. L. A study on improving close and distant device movement pose manipulation for hand-held augmented reality. **Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology - VRST '16**, p. 121–128, 2016. Available from Internet: <http://dl.acm.org/citation.cfm?doid=2993369.2993380>.

SAURO, J.; DUMAS, J. S. Comparison of three one-question, post-task usability questionnaires. In: ACM. **Proceedings of the Conference on Human Factors in Computing Systems**. [S.l.], 2009. p. 1599–1608.

SCHULTHEIS, U. et al. Comparison of a two-handed interface to a wand interface and a mouse interface for fundamental 3d tasks. In: IEEE. **3D User Interfaces (3DUI), 2012 IEEE Symposium on**. [S.l.], 2012. p. 117–124.

SHEPARD, R. N.; METZLER, J. Mental Rotation of Three-Dimensional Objects. **Science**, v. 171, n. 3972, p. 701–703, feb 1971. ISSN 0036-8075. Available from Internet: <http://www.sciencemag.org/cgi/doi/10.1126/science.171.3972.701>.

SIMEONE, A. L. Indirect touch manipulation for interaction with stereoscopic displays. **2016 IEEE Symposium on 3D User Interfaces, 3DUI 2016 - Proceedings**, p. 13–22, 2016.

SIMEONE, A. L.; GELLERSENY, H. Comparing indirect and direct touch in a stereoscopic interaction task. **2015 IEEE Symposium on 3D User Interfaces, 3DUI 2015 - Proceedings**, n. May, p. 105–108, 2015.

SMITH, S. **Digital signal processing: a practical guide for engineers and scientists**. [S.l.]: Newnes, 2013. ISBN 978-0750674447.

SONG, P. et al. WYSIWYF: Exploring and Annotating Volume Data with a Tangible Handheld Device. In: **Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11**. New York, New York, USA: ACM Press, 2011. D, p. 1333. ISBN 9781450302289. Available from Internet: <http://doi.acm.org/10.1145/1978942.1979140http://dl.acm.org/citation.cfm?doid=1978942.1979140>.

STOAKLEY, R.; CONWAY, M. J.; PAUSCH, R. Virtual reality on a WIM. **Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '95**, p.

265–272, 1995. Available from Internet: <http://portal.acm.org/citation.cfm?doid=223904.223938>.

STROTHOFF, S.; VALKOV, D.; HINRICHS, K. Triangle cursor. In: **Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces - ITS '11**. New York, New York, USA: ACM Press, 2011. p. 111. ISBN 9781450308717. Available from Internet: <http://dl.acm.org/citation.cfm?doid=2076354.2076377>.

STUERZLINGER, W.; WINGRAVE, C. A. The value of constraints for 3d user interfaces. **Virtual Realities: Dagstuhl Seminar 2008**, p. 203–223, 2011.

SUNDÉN, E.; LUNDGREN, I.; YNNERMAN, A. Hybrid Virtual Reality Touch Table - An Immersive Collaborative Platform for Public Explanatory Use of Cultural Objects and Sites. p. 109–113, 2017. ISSN 2312-6124.

ULINSKI, A. C. et al. Selection performance based on classes of bimanual actions. **3DUI - IEEE Symposium on 3D User Interfaces 2009 - Proceedings**, p. 51–58, 2009.

VEIT, M.; CAPOBIANCO, A.; BECHMANN, D. Influence of degrees of freedom's manipulation on performances during orientation tasks in virtual reality environments. **Proceeding VRST '09 Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology**, v. 1, n. 212, p. 51–58, 2009. Available from Internet: <http://portal.acm.org/citation.cfm?id=1643942>.

VINAYAK et al. MobiSweep: Exploring Spatial Design Ideation Using a Smartphone as a Hand-held Reference Plane. In: **Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '16**. New York, New York, USA: ACM Press, 2016. p. 12–20. ISBN 9781450335829. Available from Internet: <http://dl.acm.org/citation.cfm?doid=2839462.2839490>.

VUIBERT, V.; STUERZLINGER, W.; COOPERSTOCK, J. R. Evaluation of Docking Task Performance Using Mid-air Interaction Techniques. **Proceedings of the Symposium on Spatial User Interaction**, p. 44–52, 2015.

WANG, J. et al. Escape from Meadwyn 4: A cross-platform environment for collaborative navigation tasks. **IEEE Symposium on 3D User Interfaces 2012, 3DUI 2012 - Proceedings**, p. 179–180, 2012.

ZHAI, S. Human Performance in Six Degree of Freedom Input Control. **Industrial Engineering**, p. 179, 1995. Available from Internet: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.71.2326{&}rep=rep1{&}ty>.

ZHAI, S.; MILGRAM, P.; BUXTON, W. The influence of muscle groups on performance of multiple degree-of-freedom input. **Proceedings of the SIGCHI conference on Human factors in computing systems common ground - CHI '96**, p. 308–315, 1996. Available from Internet: <http://portal.acm.org/citation.cfm?doid=238386.238534>.

APPENDIX A — INTERAÇÕES 3D COLABORATIVAS E SUA APLICAÇÃO EM INTERFACES DE REALIDADE VIRTUAL, AUMENTADA E MISTA

Os humanos cooperam em uma ampla gama de atividades. Eles podem, por exemplo, organizar-se em divisões complexas para realizar tarefas que sozinhos seriam difíceis. A cooperação pode aumentar a produtividade e fornecer resultados mais rápidos (LAUGHLIN, 2011). Hoje em dia, uma quantidade considerável de trabalho é suportada por computadores e, naturalmente, precisamos evoluir as ferramentas para suportar a colaboração em tais espaços digitais. Ambientes Virtuais Colaborativos (CVEs) são espaços compartilhados projetados para suportar interações entre usuários e objetos em Realidade Virtual (VR) (CHURCHILL; SNOWDON, 1998). Devido à natureza ilimitada do virtual, além de apenas simular os aspectos do mundo físico, é possível extrapolar os limites do real para aprimorar as habilidades e percepções do usuário durante a experiência da realidade virtual. Os CVEs adicionam a possibilidade de integrar realidades mistas onde os usuários podem coexistir em uma experiência assimétrica.

Manipulações 3D virtuais são uma das tarefas canônicas em VEs (Bowman A. D; Kruijff, 2005). É a ação de alterar as propriedades de posição, rotação e escala dos elementos virtuais. Objetos virtuais são intangíveis em 3D VEs, e a interação com eles só é alcançada por meio de rastreamento de todo o corpo ou mediada por um dispositivo de interface do usuário. Criar interfaces eficazes para sistemas virtuais 3D é um desafio (ZHAI, 1995; STUERZLINGER; WINGRAVE, 2011). A capacidade de interação depende do mapeamento de entradas do usuário na ação correta. Além disso, a visualização com projeções, displays estereoscópicos e HMD oferecem diferentes níveis de imersão e afetam a experiência do usuário ao interagir com objetos virtuais. Em tarefas de manipulação 3D, é difícil ser preciso ao trabalhar de uma única perspectiva. A realização de manipulações 3D simples, como o posicionamento preciso de um objeto, requer o movimento contínuo do usuário para verificar se há partes obstruídas. Isso é tedioso e demorado, especialmente quando a forma dos objetos é complexa. Embora essas manipulações sejam obrigatórias em qualquer ambiente virtual interativo 3D, executá-las sozinha exige um alto esforço cognitivo para coordenar as possíveis transformações (VEIT; CAPOBIANCO; BECHMANN, 2009). Muitas 3D *User Interfaces*(3DUI) foram propostas para permitir interação natural, precisa e rápida com ambientes virtuais (Bowman A. D; Kruijff, 2005). As técnicas são tipicamente projetadas para aplicações específicas. Alguns deles tornaram-se relativamente comuns devido ao fácil acesso ao mercado, como interfaces de console de jogos. No entanto, até hoje, as pesquisas no campo de manipulação de objetos 3D concentram-se principalmente na interação para um usuário.

Como é interessante que várias pessoas compartilhem o mesmo espaço virtual e cooperem nele, é desejável uma interface capaz de manipular entradas de muitos usuários para o trabalho cooperativo. Uma das principais vantagens de ter uma interface colaborativa é que ela fornece aos usuários uma perspectiva de ação individual da mesma cena, que também pode incluir pontos de vista compartilhados e individuais. O trabalho em equipe envolve negociações consideráveis e, à medida que os membros da equipe variam, as estratégias da equipe e os processos de realização de tarefas também mudam. A modelagem de tais interações cria novos conceitos de colaboração comparados àqueles tipicamente baseados em um cenário de usuário único (ROBINSON, 1993).

A.1 Objetivo

Nosso objetivo é explorar o trabalho colaborativo para tarefas de manipulação 3D de objetos virtuais. Um grande problema de pesquisa de atividade colaborativa é gerenciar o acesso compartilhado ao mesmo objeto para manipulações simultâneas (MARGERY; ARNALDI; PLOUZEAU, 1999). Por exemplo, no problema de *Mover de Piano* (LENGYEL et al., 1990), onde um piano precisa ser movido por um corredor enquanto evita colisões, dois ou mais usuários precisam negociar ações, controlar forças e coordenar movimentos para levar com sucesso o piano para o destino final. Nesta tese, pesquisamos e propomos manipulações simultâneas eficientes e trabalho paralelo. Realizamos estudos guiados pelo usuário para avaliar interfaces colaborativas. Nessas avaliações, abordamos problemas de pesquisa abertos que são fundamentais para interações cooperativas contínuas: divisão de tarefas e organização de grupos, tamanhos de grupos para produtividade ideal, conscientização individual e de grupo e responsabilidade individual e de grupo por interações cooperativas em espaços virtuais 3D.

Nós criamos e aplicamos o modelo *Sum-of-Contributions* no projeto de 3DUIs colaborativas usando hardware recente e acessível: Dispositivos Móveis e *Head-Mounted Displays* (HMDs). Dispositivos móveis não são projetados para interação em VEs. No entanto, eles encapsulam várias características que os tornam eficientes ferramentas de interação 3D, como poder de processamento, sensores inerciais, conexão sem fio e telas sensíveis ao toque. Além disso, os smartphones são onipresentes na vida moderna e provavelmente serão o dispositivo de entrada mais amplamente disponível, o que é desejável ao projetar ambientes virtuais que funcionem em diferentes plataformas. Além disso, usamos dispositivos móveis como uma extensão de nosso corpo (BALLAGAS et al., 2006), e podemos supor que usuários de todas as idades e origens desenvolveram familiaridade com as interfaces de entrada de dispositivos móveis, o que poderia resultar em uma curva de aprendizado relativamente curta seus recursos são explorados no contexto da 3DUI. Por outro lado, os HMDs são concebidos para experiências imersivas em 3D.

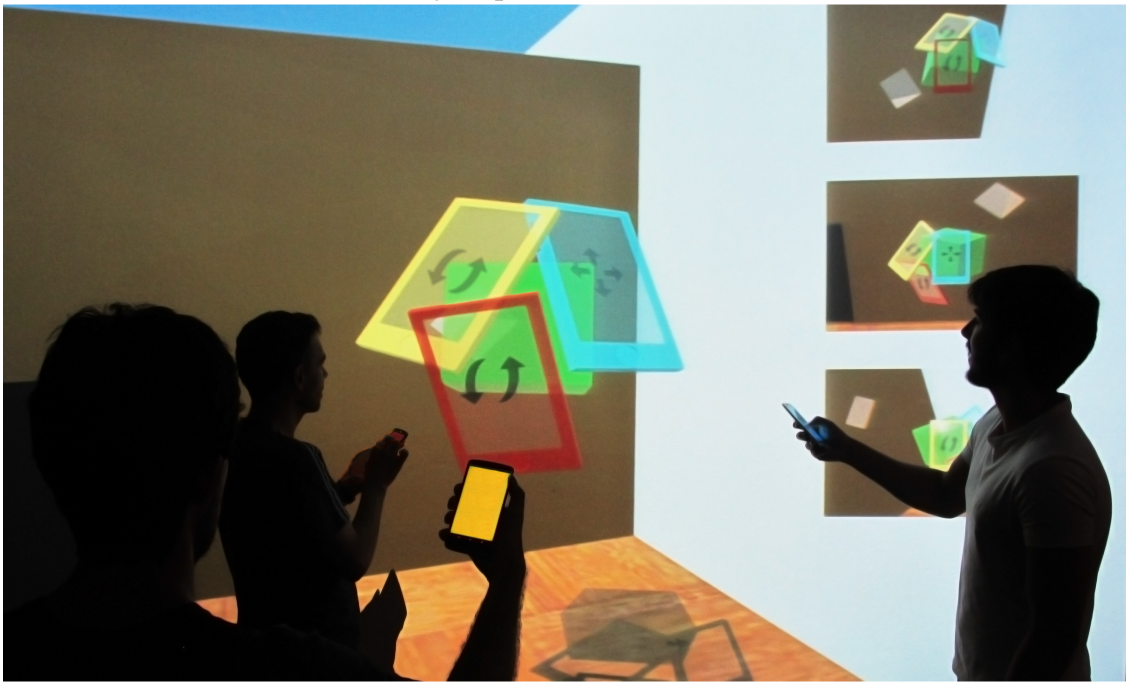
No final, fornecemos soluções colaborativas para as principais plataformas de exibição, envolvendo realidade aumentada e virtual. Além disso, soluções individuais são compatíveis entre si. Desta forma, eles podem ser misturados em várias configurações para se adaptar às necessidades da tarefa. Um número ilimitado de usuários com interfaces diferentes poderia trabalhar em conjunto. Por exemplo, alguns colaboradores poderiam trabalhar imersos em HMD, enquanto outros poderiam visualizar a tarefa em uma tela ou com dispositivos de realidade aumentada. Esta pesquisa também contribui para a definição de linhas de base para o design de interações colaborativas em VEs.

A.2 Interface 3D baseada em dispositivos portáteis para manipulação colaborativa de objetos em ambientes virtuais

Nessa primeira parte, apresentamos o projeto de uma nova interface de usuário 3D para manipulação colaborativa de objetos em ambientes virtuais 3D, a fim de avaliar a cooperação em grupos (ver Figura A.1). A técnica é baseada em smartphones e usa a tela sensível ao toque e os sensores inerciais como uma 3DUI. A técnica provou ser intuitiva e robusta. Mais de cem participantes testaram e puderam fornecer dados e feedback sobre sua usabilidade. Em vários testes informais, os usuários foram convidados a baixar o aplicativo em seus próprios *smartphones* e participar de uma sessão de manipulação em andamento. Nós testamos com equipes compostas por mais de dez membros usando smartphones de diferentes modelos. Todos eles entenderam com facilidade e rapidez a finalidade e a mecânica da técnica, indicando uma alta disponibilidade.

Em experimentos formais com usuários, demonstramos que equipes com mais membros executavam com mais precisão do que grupos menores. Em grupos maiores, os membros poderiam dividir melhor as tarefas básicas, assumindo papéis e, como conseqüência, diminuindo a carga de trabalho. A velocidade não foi o objetivo em nossos experimentos, mas também observamos que a velocidade aumenta desde o primeiro uso do sistema até o segundo sem afetar a precisão, demonstrando que os usuários aprendem

Figure A.1: Três usuários simultaneamente dirigindo o objeto pelo ambiente virtual. Os retângulos coloridos indicam a posição e orientação de cada usuário no VE. As três janelas à direita mostram as três visualizações pessoais.

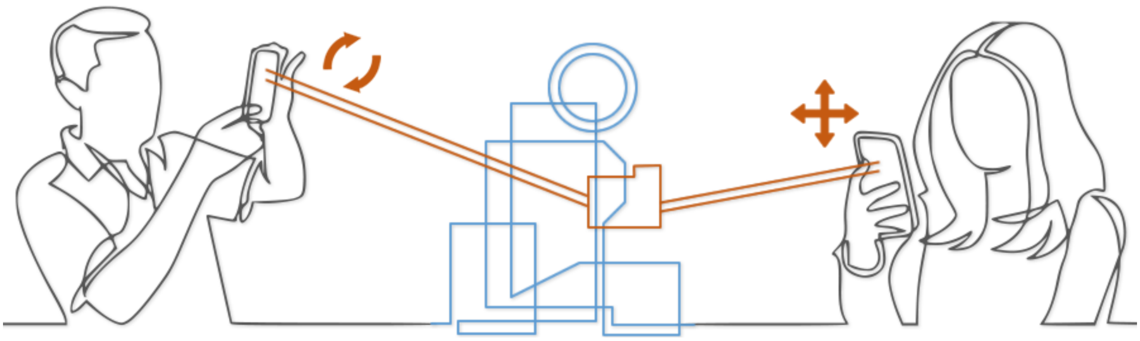


Fonte: o Autor

rápido. No entanto, mais testes devem ser feitos em relação à curva de aprendizado da técnica. Nossos resultados indicam que em tarefas de manipulação virtual, onde a precisão é obrigatória, compartilhar os aspectos de manipulação aumenta significativamente a precisão.

A.3 Técnica de interação 3D colaborativa para realidade aumentada portátil

Figure A.2: Dois usuários manipulando simultaneamente um objeto virtual em realidade aumentada. Cada usuário pode ter uma visão em perspectiva diferente da cena. Os raios são desenhados entre o dispositivo e o objeto selecionado informam aos usuários a seleção. Ícones virtuais indicam a transformação atual.



Fonte: o Autor

Com base na experiência adquirida no experimento anterior, neste capítulo, apresentamos o design de uma nova interface de usuário para manipulação colaborativa de objetos virtuais 3D sobrepostas ao ambiente físico. A técnica foi projetada para dispositivos portáteis, pois são versáteis e onipresentes nas tarefas diárias. Projetamos dois modos para manipulações 3D, gestos de toque e movimentos de dispositivos, que combinados permitem transformações intuitivas de 7-DOF. Graças ao modelo *Sum-of-Contributions*, a técnica cria um meio compartilhado onde vários usuários podem interagir simultaneamente com seus próprios dispositivos. A técnica resolve o problema de manipulações simultâneas com uma abordagem de coordenação de ação, onde cada contribuição de cada usuário conta como uma etapa de transformação e é aplicada diretamente ao objeto.

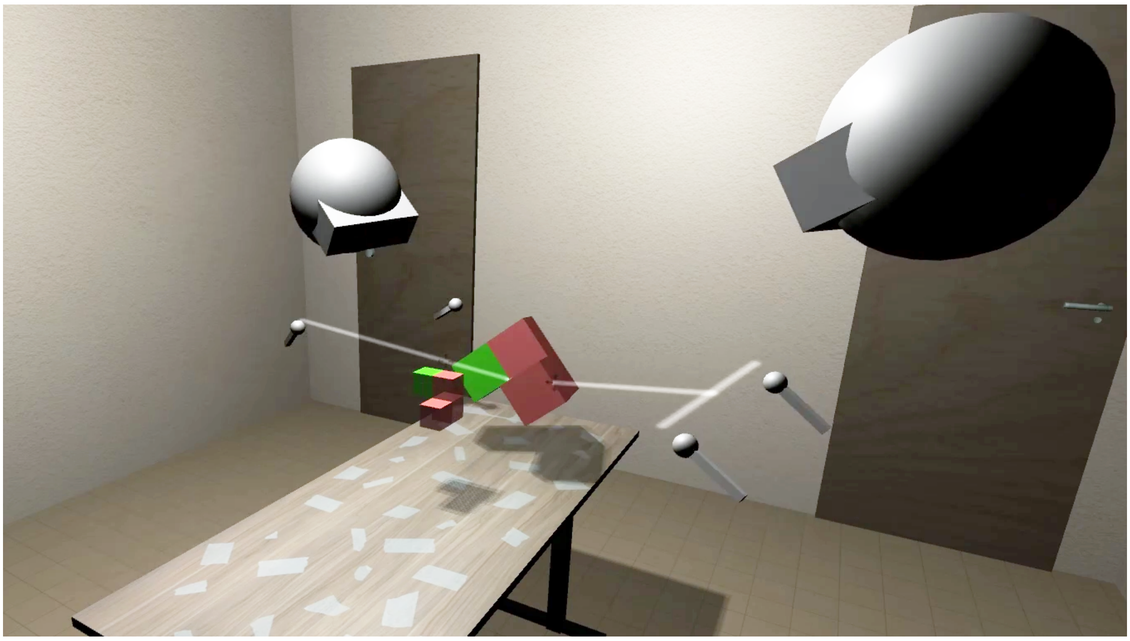
Mais de cinquenta participantes testaram a interface em demonstrações públicas. Eles poderiam baixar o aplicativo em seus smartphones e participar de demonstrações em andamento. Até cinco membros da equipe usando dispositivos de diferentes modelos foram registrados durante as demonstrações. Todos eles entenderam com facilidade e rapidez a finalidade e a mecânica da técnica, indicando uma alta disponibilidade. A observação de grupos durante as demonstrações inspirou o design dos dois experimentos.

Demonstramos a eficácia da abordagem *Hybrid* quando comparada com os métodos *Touch Gestures* ou *Movements* na avaliação da interface. Além disso, observamos que os usuários podem alternar facilmente entre os métodos e usar a ação mais eficiente para transformar corretamente o objeto e, ao mesmo tempo, manter um desempenho de alta qualidade. No segundo experimento com pares, observamos que duas estratégias foram adotadas quando não impusemos nenhuma restrição à colaboração. Curiosamente, os pares com estratégias diferentes obtiveram desempenho de tarefa semelhante, enquanto as equipes que privilegiaram a estratégia *Interação Compartilhada* se sentiram mais envolvidas na tarefa.

A.4 Manipulações 3D colaborativas para realidade virtual imersiva

Neste capítulo, estendemos os conceitos de interações colaborativas para a RV imersiva (Figura A.3). Primeiro, projetamos uma interface de usuário capaz de lidar com manipulações simultâneas no mesmo objeto virtual. A técnica foi projetada para HMDs e seus controles, já que são o hardware mais intuitivo e imersivo disponível no momento. Provamos a versatilidade do nosso modelo *Sum-of-Contributions* no desenvolvimento de uma técnica de interação para manipulação de objetos à distância. Nossa abordagem supera

Figure A.3: Dois usuários interagindo simultaneamente com um objeto 3D em realidade virtual imersiva. A técnica permite manipulações além do alcance do braço. Raios desenhados da mão para o objeto representam a seleção. Os usuários podem optar por interagir com uma ou duas mãos. Formas simples representam os avatares do usuário no VE.



Source: the Author

a limitação egocêntrica do *Virtual Hand* em troca de um desempenho geral mais baixo no tempo. As vantagens do nosso design tornam-se evidentes no trabalho colaborativo. Embora o desempenho em pares seja mais baixo do que em indivíduos, a carga de trabalho significativamente menor indica que os usuários colaboram de forma que a quantidade de trabalho seja distribuída entre os participantes.

A.5 Colaboração Assimétrica com Técnicas 3DUI Mistas

Neste capítulo, propusemos uma abordagem assimétrica para co-manipulação em um ambiente compartilhado. Os usuários de RV e RA coexistem no mesmo espaço físico, mantendo sua perspectiva de vista, escala e dimensionalidade (ver Figura 6.1). Junto com o projeto de abordagem, realizamos um estudo de usuário para comparar interações assimétricas e simétricas em um VE colaborativo. Nós variamos a realidade que os usuários experimentam para avaliar o desempenho, participação e presença social. As interfaces de usuário selecionadas foram o AR e o VR propostos nos Capítulos anteriores. O experimento foi conduzido em pares em duas condições simétricas, VR-VR e AR-AR e uma condição assimétrica VR-AR. Nós demonstramos que os pares na assimetria VR-AR alcançaram um

Figure A.4: Interação simultânea assimétrica com técnicas 3DUI mistas. Dois usuários co-localizados estão interagindo simultaneamente com um objeto 3D usando diferentes interfaces enquanto compartilham o mesmo espaço físico.

Source: the Author

desempenho melhor do que a condição AR-AR, e um desempenho um pouco pior do que o VR-VR. Em todas as situações, os participantes tiveram uma participação de trabalho similar quando manipulavam os mesmos objetos ou objetos diferentes. Isso indica que os pares estão cooperando para resolver a tarefa. Embora tenhamos relatado menor assistência mútua na configuração assimétrica, observamos que os participantes frequentemente falam sobre suas ações e estratégias durante o experimento. Essa informação foi comprovada por questionários pós-experimento, onde vários usuários relataram uma estratégia de divisão do trabalho.

A.6 Conclusão

Nesta tese, apresentamos uma pesquisa que focou em manipulações simultâneas de objetos 3D em ambientes virtuais colaborativos. Visamos o *design* de um modelo robusto capaz de manipular entradas de várias fontes, que chamamos de *Sum-of-Contributions* (SoC). Aplicamos o modelo SoC no design de interfaces de usuário 3D para colaboração simétrica e assimétrica em cenários virtuais e aumentados. Levamos em conta as características de hardware de entrada, os recursos de exibição e as necessidades de reconhecimento durante o processo de design. Nossas soluções foram baseadas na integração e adaptação das melhores práticas presentes na literatura. Junto com o design da interface, avaliamos o comportamento dos usuários durante tarefas colaborativas. Abaixo resumimos as fases relatadas nesta tese:

1. Nós exploramos o uso de dispositivos portáteis como um controlador para uma segunda tela. Nesse cenário, um grupo de usuários co-manipulou um objeto virtual ao compartilhar a mesma tela e ponto de vista. Nós investigamos a relação entre os tamanhos dos grupos e o tempo e a precisão para concluir as tarefas de manipulação;

2. Em uma segunda fase, estendemos a interface para manipulações colaborativas em ambientes de realidade aumentada. Aqui, o tamanho de tela limitado é disputado pela visualização e interação, representando novos desafios. Como benefício, os usuários têm seu ponto de vista personalizado dependendo de sua posição no ambiente. Nós avaliamos três projetos de interação e avaliamos a capacidade de pares para resolver a tarefa de encaixe, permitindo que eles escolhessem suas estratégias de colaboração;

3. Nós aplicamos o modelo SoC para interações colaborativas de RV imersivas. Primeiro, comparamos a abordagem *Virtual Hand*, que é a interação mais natural para interação de RV imersiva, com nossa *Interação à Distância* baseada na abordagem *HOMER* em uma avaliação de usuário. Em seguida, comparamos o desempenho e a carga de trabalho do trabalho de pares de usuários com relação a usuários individuais;

4. No final, integramos as soluções AR e VR em um ambiente colaborativo assimétrico. Comparamos os aspectos de desempenho e colaboração da abordagem assimétrica com dois projetos simétricos, onde os pares interagiam apenas na realidade aumentada e apenas na realidade virtual.

Nós encerramos com um conjunto de considerações de *design* e *insights* sobre os pontos essenciais relatados nesta tese.

A colaboração pode ser benéfica em tarefas 3D que a precisão é necessária. Se a tarefa requer velocidade acima da precisão, é necessário levar em consideração as interfaces envolvidas e como os usuários são integrados ao ambiente para resolver a tarefa. Em nossos estudos, não conseguimos encontrar nenhuma vantagem na velocidade de colaboração sobre usuários únicos. Em relação ao número de participantes manipulando simultaneamente o mesmo objeto, observamos que adicionar mais elementos em um grupo pode causar uma queda no desempenho individual (efeito de Ringelmann). Embora sejam necessários mais estudos sobre o tamanho dos grupos, recomendamos manter o tamanho do grupo com o máximo de quatro participantes manipulando o mesmo objeto. Também notamos que a participação na tarefa é balanceada entre os usuários em pares. Assim, espera-se que, se os pares forem bem treinados, ambos funcionem igualmente, mesmo em configurações assimétricas.

Em todos os estudos, observamos que os usuários planejam estratégias para resolver as tarefas de manipulação, mesmo que os participantes não se conheçam. Os participantes não relataram nenhum desconforto ao trabalhar com estranhos. Se as manipulações simultâneas são obrigatórias, descobrimos que a estratégia de divisão do trabalho resultou no melhor desempenho do que os grupos com outras estratégias. Em situações em que os pares podem escolher se manipulam os mesmos objetos ou objetos diferentes para resolver uma tarefa, observamos que ambas as abordagens têm desempenho semelhante. No entanto, os pares que trabalharam juntos no mesmo objeto se sentiram mais envolvidos durante as tarefas.

Com exceção das interfaces de usuário tangíveis (UISs), a abordagem de interação indireta é mais adequada para manipulações colaborativas. Mesmo que a abordagem direta, usando a técnica da mão virtual, por exemplo, seja a interação mais natural, ela é bastante egocêntrica e não é amigável para manipulações simultâneas. Por outro lado, as manipulações em 3D à distância, embora menos naturais e mais lentas que uma abordagem direta, podem ser implementadas aproveitando as entradas simultâneas de diferentes fontes ao mesmo tempo, mantendo boa disponibilidade.

Destacamos que fornecer feedback visual das seleções e ações é essencial para a conscientização do usuário durante a tarefa colaborativa, mesmo na colaboração em ambiente físico compartilhado, onde a comunicação não era um problema. A capacidade de mostrar o que estão fazendo melhora o conhecimento sobre a tarefa e os outros.

APPENDIX B — LIST OF PUBLICATIONS AND AWARDS

International Journals:

- **Assessment of Asymmetric Collaborative Interactions in Virtual and Augmented Realities**
  Jerônimo G. Grandi, Henrique G. Debarba and Anderson Maciel.
  IEEE Transactions on Visualization and Computer Graphics
  Submitted

International Conferences:

- **Design and Assessment of a Collaborative 3D Interaction Technique for Handheld Augmented Reality**
  Jerônimo G. Grandi, Henrique G. Debarba, Iago Berndt, Luciana Nedel and Anderson Maciel.
  IEEE VR Conference on Virtual Reality and 3D User Interfaces (2018)
  Honorable Mention for Best Conference Paper
  DOI: https://doi.org/10.1109/VR.2018.8446295

- **Design and Evaluation of a Handheld-based 3D User Interface for Collaborative Object Manipulation**
  Jerônimo G. Grandi, Henrique G. Debarba, Luciana Nedel and Anderson Maciel.
  ACM CHI Conference on Human Factors in Computing Systems (2017)
  DOI: http://dx.doi.org/10.1145/3025453.3025935

Demonstrations Papers:

- **3DAthlon: 3D Gestural Interfaces to Support a 3-Stage Contest in VR**
  Jerônimo G. Grandi, Henrique G. Debarba, Juliano Franz, Victor Oliveira, Abel Ticona, Gabrielle Souza, Izadora Berti, Steeven Villa, Luciana Nedel and Anderson Maciel.
  IEEE VR Conference on Virtual Reality and 3D User Interfaces (2018)
  DOI: https://doi.org/10.1109/VR.2018.8446286

- **Collaborative Manipulation of 3D Virtual Objects in Augmented Reality Scenarios using Mobile Devices**

Jerônimo G. Grandi, Iago Berndt, Henrique G. Debarba, Luciana Nedel and Anderson Maciel.
IEEE Symposium on 3D User Interfaces (2017)
Best 3DUI Contest Award
DOI: https://doi.org/10.1109/3DUI.2017.7893373

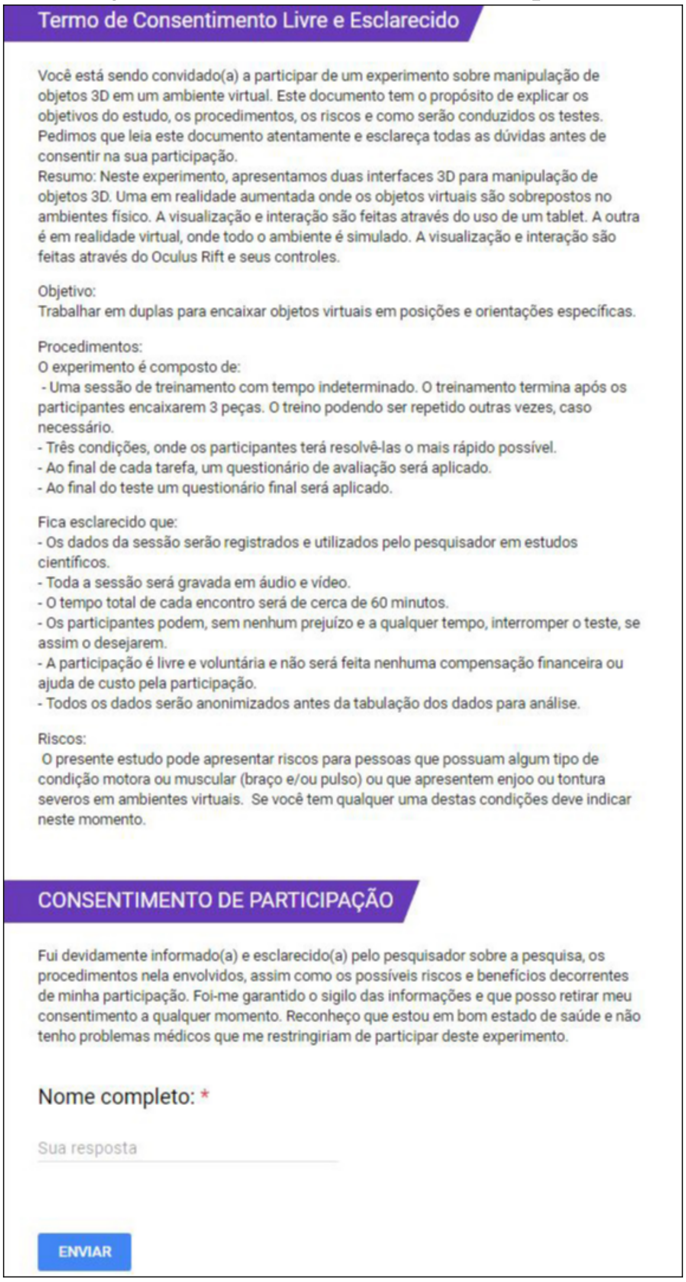
- **Collaborative 3D Manipulation using Mobile Phones**
  Jerônimo G. Grandi, Iago Berndt, Henrique G. Debarba, Luciana Nedel and Anderson Maciel.
  IEEE Symposium on 3D User Interfaces (2016)
  Best 3DUI Contest Award
  DOI: http://dx.doi.org/10.1109/3DUI.2016.7460079

- **Handheld-based 3D User Interface for Collaborative Object Manipulation**
  Jerônimo G. Grandi and Anderson Maciel.
  ACM CHI Conference on Human Factors in Computing Systems (2017)

- **A Collaborative 3D Manipulation Challenge**
  Jerônimo G. Grandi and Anderson Maciel.
  Symposium on Virtual and Augmented Reality (SVR) (2016)

APPENDIX C — CONSENT INFORM FORM

Figure C.1: Consent Inform Form (pt-br)



Termo de Consentimento Livre e Esclarecido

Você está sendo convidado(a) a participar de um experimento sobre manipulação de objetos 3D em um ambiente virtual. Este documento tem o propósito de explicar os objetivos do estudo, os procedimentos, os riscos e como serão conduzidos os testes. Pedimos que leia este documento atentamente e esclareça todas as dúvidas antes de consentir na sua participação.
Resumo: Neste experimento, apresentamos duas interfaces 3D para manipulação de objetos 3D. Uma em realidade aumentada onde os objetos virtuais são sobrepostos no ambientes físico. A visualização e interação são feitas através do uso de um tablet. A outra é em realidade virtual, onde todo o ambiente é simulado. A visualização e interação são feitas através do Oculus Rift e seus controles.

Objetivo:
Trabalhar em duplas para encaixar objetos virtuais em posições e orientações específicas.

Procedimentos:
O experimento é composto de:
- Uma sessão de treinamento com tempo indeterminado. O treinamento termina após os participantes encaixarem 3 peças. O treino podendo ser repetido outras vezes, caso necessário.
- Três condições, onde os participantes terá resolvê-las o mais rápido possível.
- Ao final de cada tarefa, um questionário de avaliação será aplicado.
- Ao final do teste um questionário final será aplicado.

Fica esclarecido que:
- Os dados da sessão serão registrados e utilizados pelo pesquisador em estudos científicos.
- Toda a sessão será gravada em áudio e vídeo.
- O tempo total de cada encontro será de cerca de 60 minutos.
- Os participantes podem, sem nenhum prejuízo e a qualquer tempo, interromper o teste, se assim o desejarem.
- A participação é livre e voluntária e não será feita nenhuma compensação financeira ou ajuda de custo pela participação.
- Todos os dados serão anonimizados antes da tabulação dos dados para análise.

Riscos:
O presente estudo pode apresentar riscos para pessoas que possuam algum tipo de condição motora ou muscular (braço e/ou pulso) ou que apresentem enjoo ou tontura severos em ambientes virtuais. Se você tem qualquer uma destas condições deve indicar neste momento.

CONSENTIMENTO DE PARTICIPAÇÃO

Fui devidamente informado(a) e esclarecido(a) pelo pesquisador sobre a pesquisa, os procedimentos nela envolvidos, assim como os possíveis riscos e benefícios decorrentes de minha participação. Foi-me garantido o sigilo das informações e que posso retirar meu consentimento a qualquer momento. Reconheço que estou em bom estado de saúde e não tenho problemas médicos que me restringiriam de participar deste experimento.

Nome completo: *

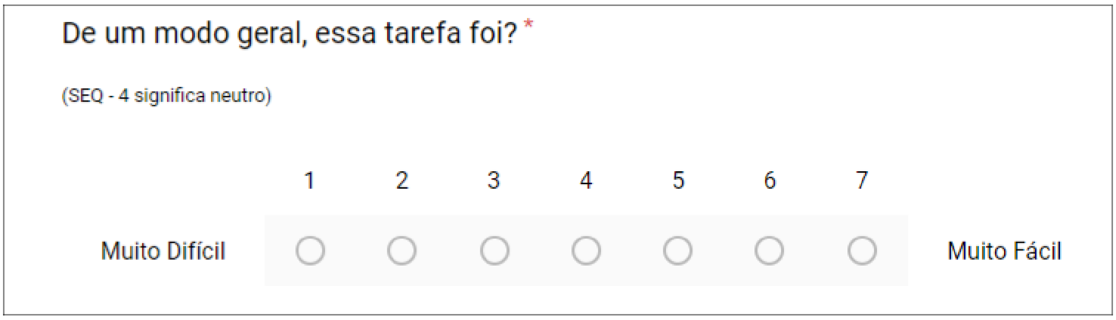
Sua resposta

ENVIAR

Source: the Author

APPENDIX D — POST-TASK USABILITY QUESTIONNAIRE

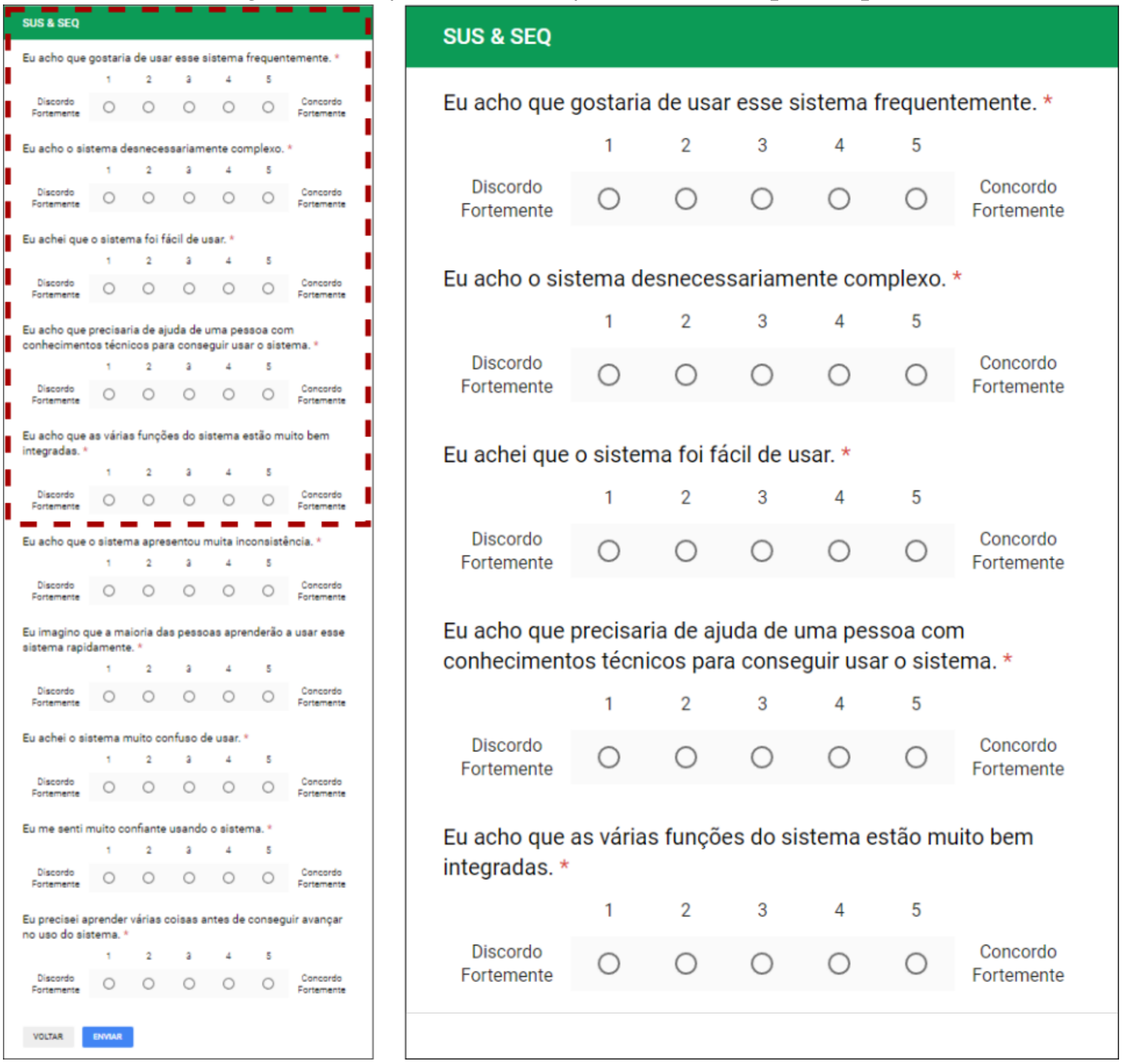
Figure D.1: Single Easy Question (pt-br)



Source: adapted from Sauro and Dumas (2009)

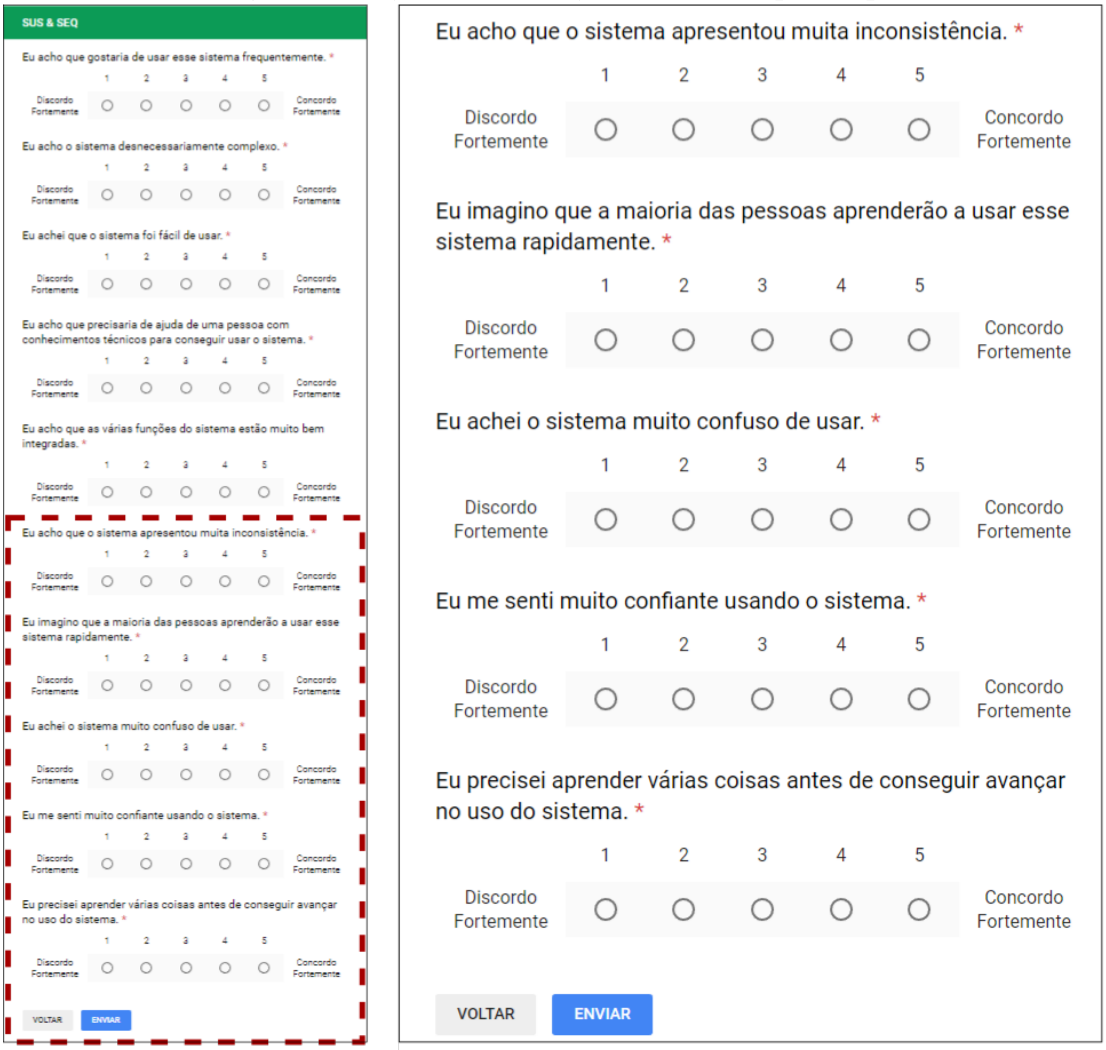
APPENDIX E — SYSTEM USABILITY QUESTIONNAIRE

Figure E.1: System Usability Scale (SUS) (pt-br) (p1)



Source: adapted from Brooke (2013)

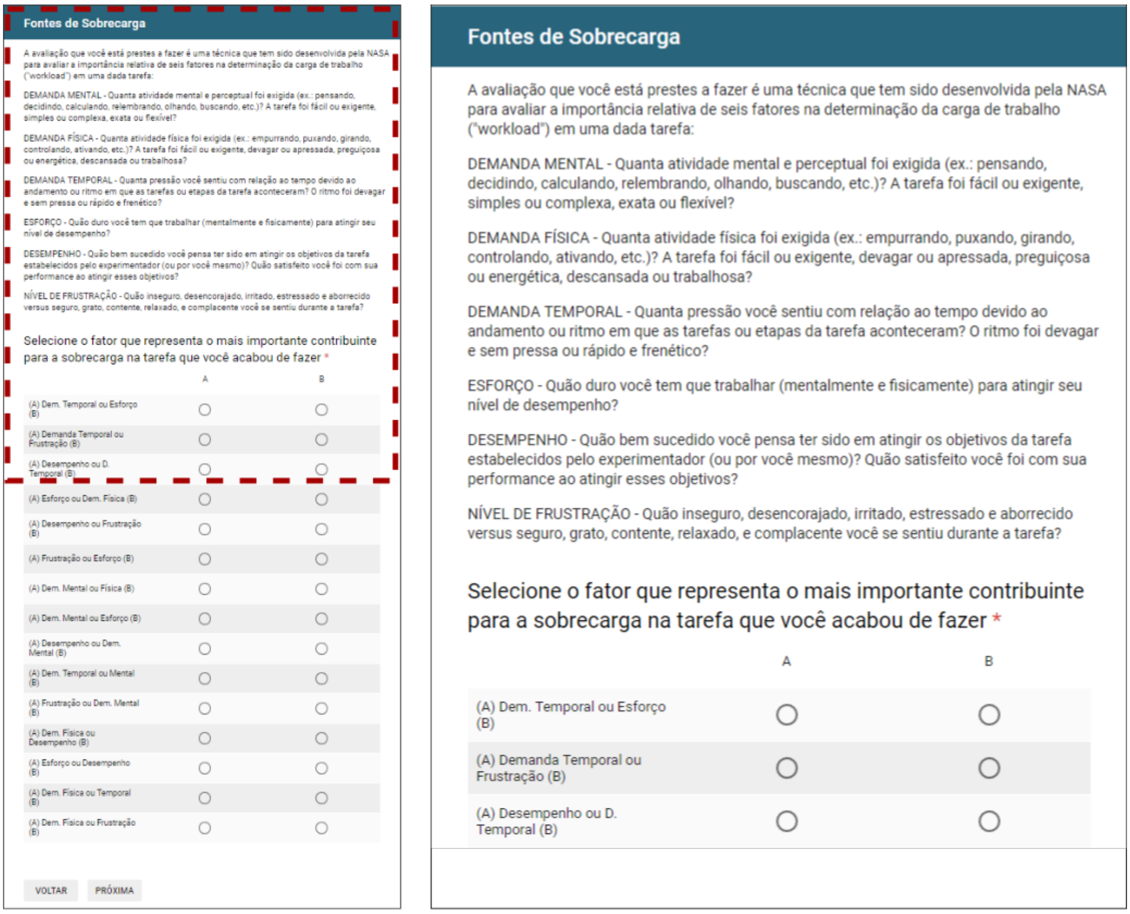
Figure E.2: System Usability Scale (SUS) (pt-br) (p2)



Source: adapted from Brooke (2013)

APPENDIX F — WORKLOAD QUESTIONNAIRE

Figure F.1: NASA TLX - Sources of Load (pt-br) (p1)



Fontes de Sobrecarga

A avaliação que você está prestes a fazer é uma técnica que tem sido desenvolvida pela NASA para avaliar a importância relativa de seis fatores na determinação da carga de trabalho ("workload") em uma dada tarefa:

DEMANDA MENTAL - Quanta atividade mental e perceptual foi exigida (ex.: pensando, decidindo, calculando, relembrando, olhando, buscando, etc.)? A tarefa foi fácil ou exigente, simples ou complexa, exata ou flexível?

DEMANDA FÍSICA - Quanta atividade física foi exigida (ex.: empurrando, puxando, girando, controlando, ativando, etc.)? A tarefa foi fácil ou exigente, devagar ou apressada, preguiçosa ou energética, descansada ou trabalhosa?

DEMANDA TEMPORAL - Quanta pressão você sentiu com relação ao tempo devido ao andamento ou ritmo em que as tarefas ou etapas da tarefa aconteceram? O ritmo foi devagar e sem pressa ou rápido e frenético?

ESFORÇO - Quão duro você tem que trabalhar (mentalmente e fisicamente) para atingir seu nível de desempenho?

DESEMPENHO - Quão bem sucedido você pensa ter sido em atingir os objetivos da tarefa estabelecidos pelo experimentador (ou por você mesmo)? Quão satisfeito você foi com sua performance ao atingir esses objetivos?

NÍVEL DE FRUSTRAÇÃO - Quão inseguro, desencorajado, irritado, estressado e aborrecido versus seguro, grato, contente, relaxado, e complacente você se sentiu durante a tarefa?

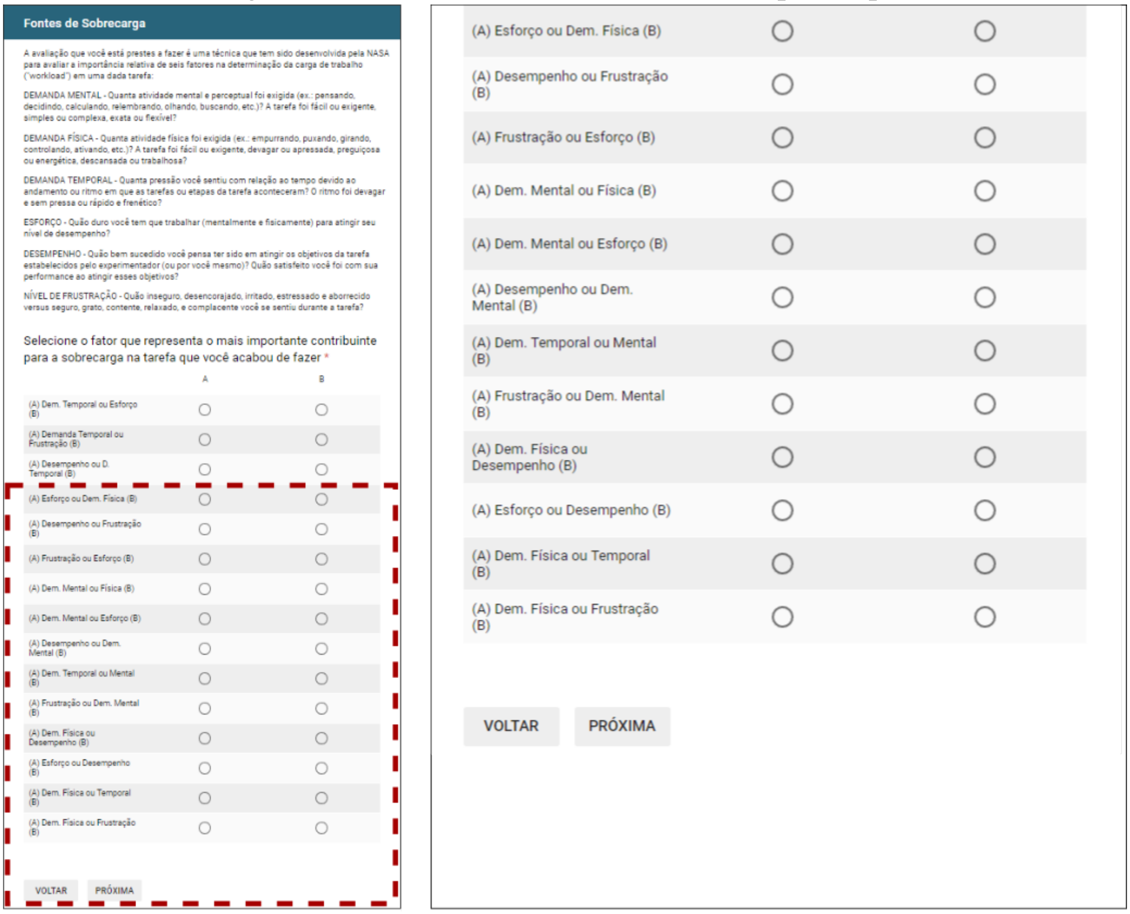
Selecione o fator que representa o mais importante contribuinte para a sobrecarga na tarefa que você acabou de fazer *

	A	B
(A) Dem. Temporal ou Esforço (B)	○	○
(A) Demanda Temporal ou Frustração (B)	○	○
(A) Desempenho ou D. Temporal (B)	○	○
(A) Esforço ou Dem. Física (B)	○	○
(A) Desempenho ou Frustração (B)	○	○
(A) Frustração ou Esforço (B)	○	○
(A) Dem. Mental ou Física (B)	○	○
(A) Dem. Mental ou Esforço (B)	○	○
(A) Desempenho ou Dem. Mental (B)	○	○
(A) Dem. Temporal ou Mental (B)	○	○
(A) Frustração ou Dem. Mental (B)	○	○
(A) Dem. Física ou Desempenho (B)	○	○
(A) Esforço ou Desempenho (B)	○	○
(A) Dem. Física ou Temporal (B)	○	○
(A) Dem. Física ou Frustração (B)	○	○

VOLTAR PRÓXIMA

Source: adapted from Hart (2006)

Figure F.2: NASA TLX - Sources of Load (pt-br) (p2)



Fontes de Sobrecarga

A avaliação que você está prestes a fazer é uma técnica que tem sido desenvolvida pela NASA para avaliar a importância relativa de seis fatores na determinação da carga de trabalho ("workload") em uma dada tarefa:

DEMANDA MENTAL - Quanta atividade mental e perceptual foi exigida (ex.: pensando, decidindo, calculando, relembrando, olhando, buscando, etc.)? A tarefa foi fácil ou exigente, simples ou complexa, exata ou flexível?

DEMANDA FÍSICA - Quanta atividade física foi exigida (ex.: empurrando, puxando, girando, controlando, ativando, etc.)? A tarefa foi fácil ou exigente, devagar ou apressada, preguiçosa ou energética, descansada ou trabalhosa?

DEMANDA TEMPORAL - Quanta pressão você sentiu com relação ao tempo devido ao andamento ou ritmo em que as tarefas ou etapas da tarefa aconteceram? O ritmo foi devagar e sem pressa ou rápido e frenético?

ESFORÇO - Quão duro você tem que trabalhar (mentalmente e fisicamente) para atingir seu nível de desempenho?

DESEMPENHO - Quão bem sucedido você pensa ter sido em atingir os objetivos da tarefa estabelecidos pelo experimentador (ou por você mesmo)? Quão satisfeito você foi com sua performance ao atingir esses objetivos?

NÍVEL DE FRUSTRAÇÃO - Quão inseguro, desencorajado, irritado, estressado e aborrecido versus seguro, grato, contente, relaxado, e complacente você se sentiu durante a tarefa?

Selecione o fator que representa o mais importante contribuinte para a sobrecarga na tarefa que você acabou de fazer *

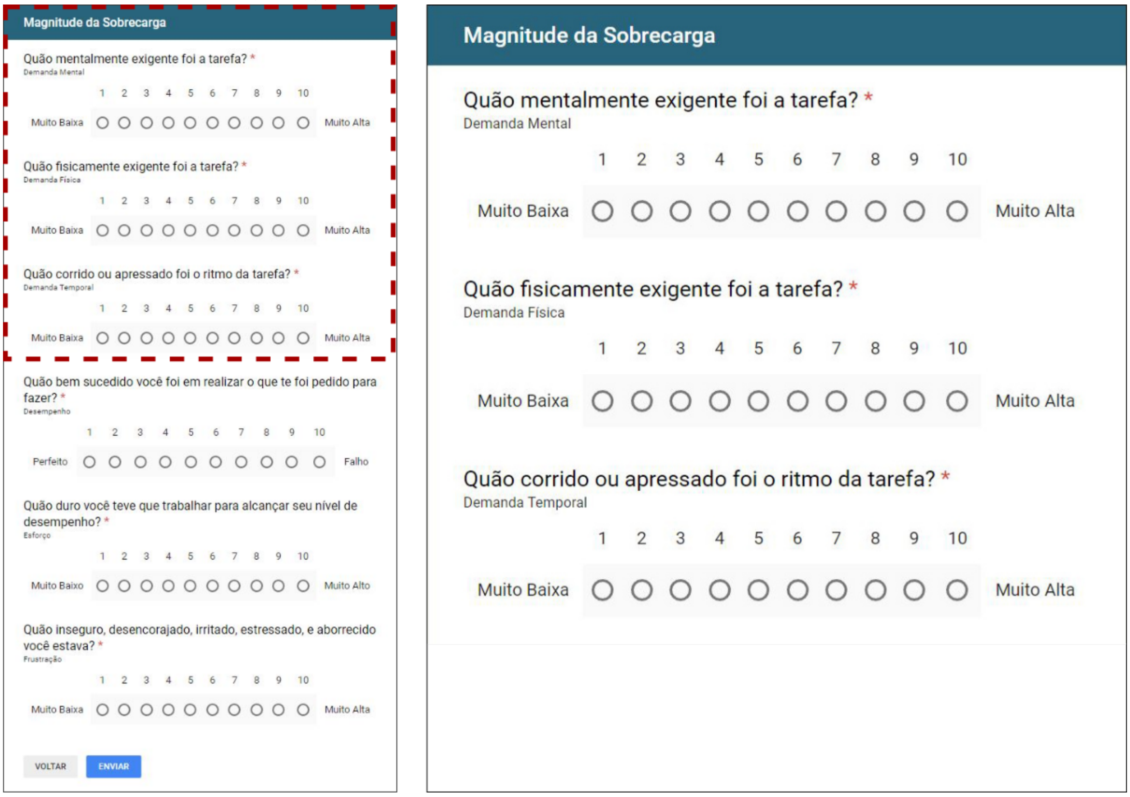
	A	B
(A) Dem. Temporal ou Esforço (B)	○	○
(A) Demanda Temporal ou Frustração (B)	○	○
(A) Desempenho ou D. Temporal (B)	○	○
(A) Esforço ou Dem. Física (B)	○	○
(A) Desempenho ou Frustração (B)	○	○
(A) Frustração ou Esforço (B)	○	○
(A) Dem. Mental ou Física (B)	○	○
(A) Dem. Mental ou Esforço (B)	○	○
(A) Desempenho ou Dem. Mental (B)	○	○
(A) Dem. Temporal ou Mental (B)	○	○
(A) Frustração ou Dem. Mental (B)	○	○
(A) Dem. Física ou Desempenho (B)	○	○
(A) Esforço ou Desempenho (B)	○	○
(A) Dem. Física ou Temporal (B)	○	○
(A) Dem. Física ou Frustração (B)	○	○

VOLTAR PRÓXIMA

(A) Esforço ou Dem. Física (B)
(A) Desempenho ou Frustração (B)
(A) Frustração ou Esforço (B)
(A) Dem. Mental ou Física (B)
(A) Dem. Mental ou Esforço (B)
(A) Desempenho ou Dem. Mental (B)
(A) Dem. Temporal ou Mental (B)
(A) Frustração ou Dem. Mental (B)
(A) Dem. Física ou Desempenho (B)
(A) Esforço ou Desempenho (B)
(A) Dem. Física ou Temporal (B)
(A) Dem. Física ou Frustração (B)
VOLTAR PRÓXIMA

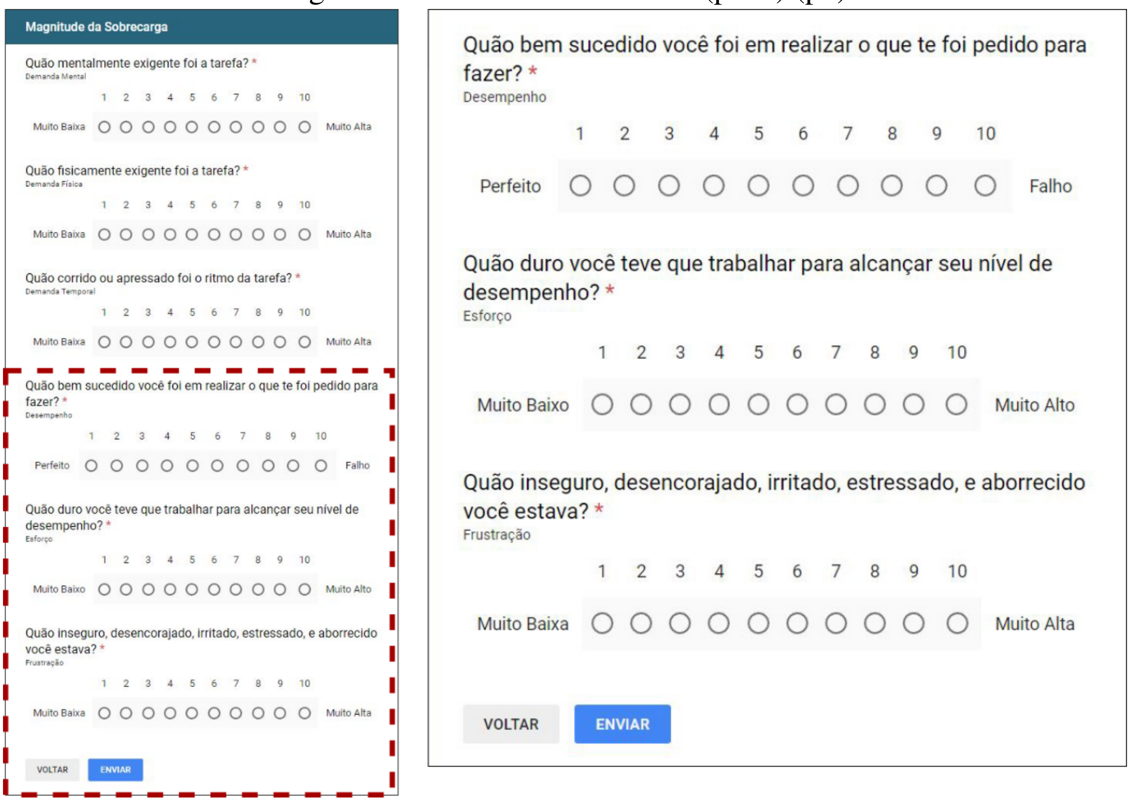
Source: adapted from Hart (2006)

Figure F.3: NASA TLX - Scales (pt-br) (p1)



Source: adapted from Hart (2006)

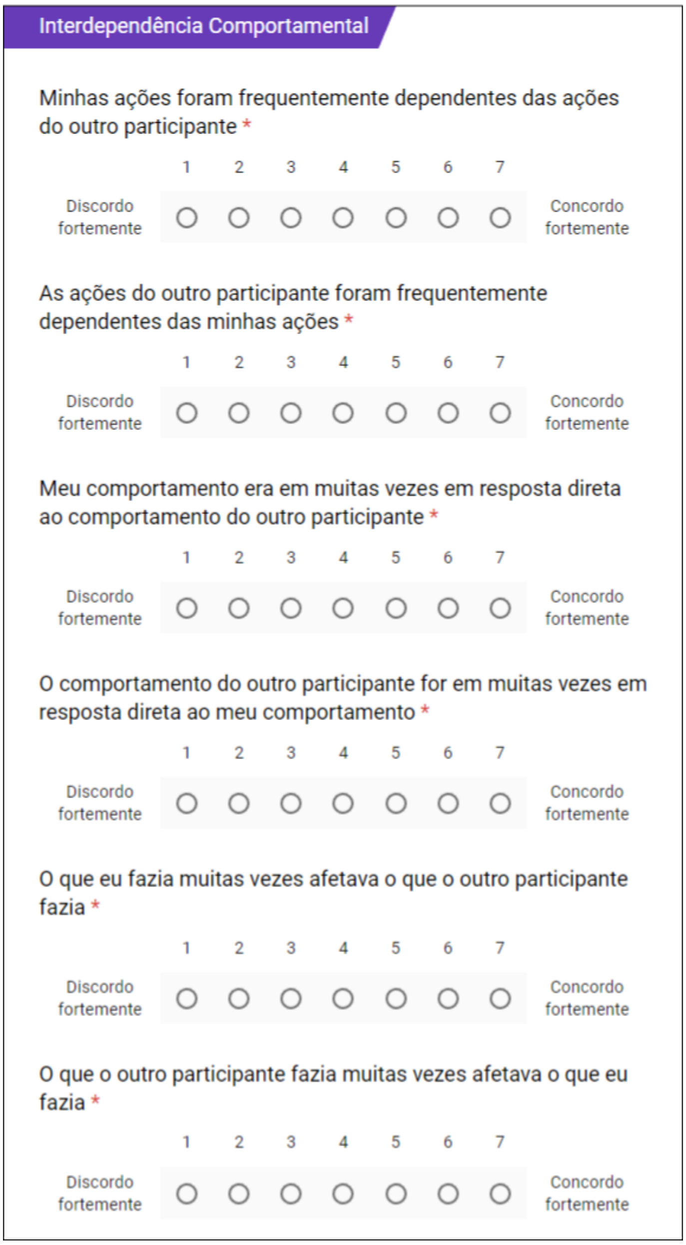
Figure F.4: NASA TLX - Scales (pt-br) (p2)



Source: adapted from Hart (2006)

APPENDIX G — SOCIAL PRESENCE QUESTIONNAIRE

Figure G.1: Networked Minds Measure of Social Presence (pt-br) (p1)



Interdependência Comportamental

Minhas ações foram frequentemente dependentes das ações do outro participante *

	1	2	3	4	5	6	7	
Discordo fortemente	○	○	○	○	○	○	○	Concordo fortemente

As ações do outro participante foram frequentemente dependentes das minhas ações *

	1	2	3	4	5	6	7	
Discordo fortemente	○	○	○	○	○	○	○	Concordo fortemente

Meu comportamento era em muitas vezes em resposta direta ao comportamento do outro participante *

	1	2	3	4	5	6	7	
Discordo fortemente	○	○	○	○	○	○	○	Concordo fortemente

O comportamento do outro participante for em muitas vezes em resposta direta ao meu comportamento *

	1	2	3	4	5	6	7	
Discordo fortemente	○	○	○	○	○	○	○	Concordo fortemente

O que eu fazia muitas vezes afetava o que o outro participante fazia *

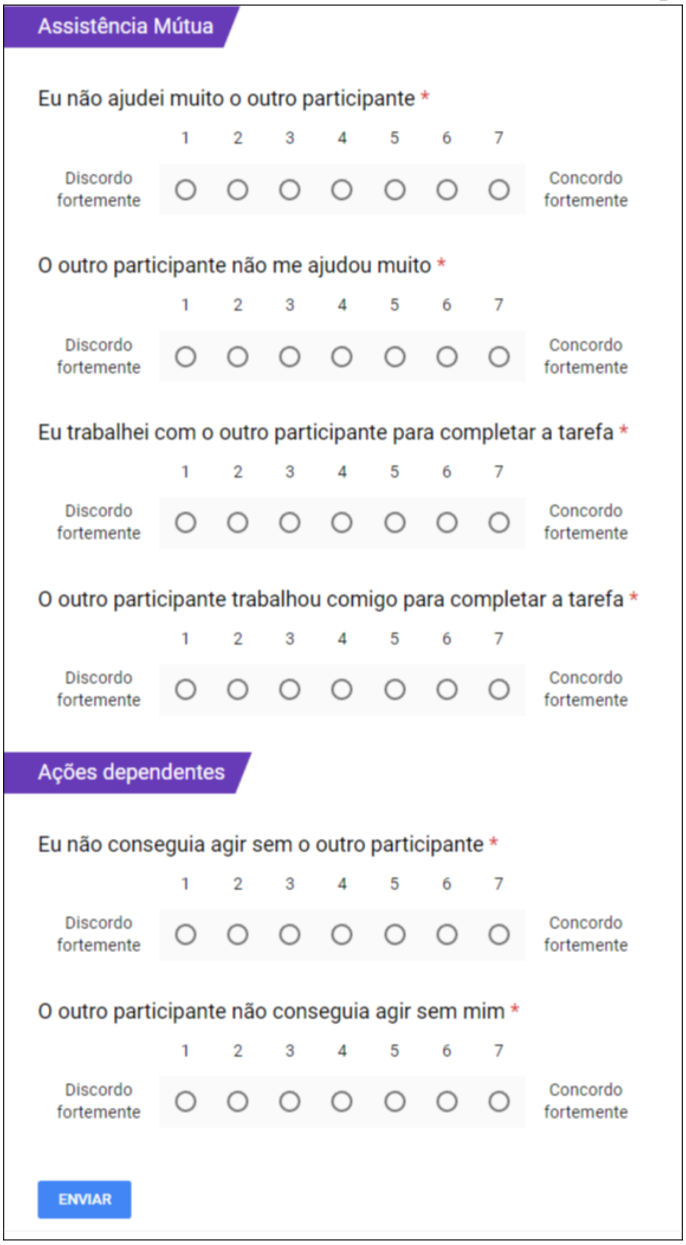
	1	2	3	4	5	6	7	
Discordo fortemente	○	○	○	○	○	○	○	Concordo fortemente

O que o outro participante fazia muitas vezes afetava o que eu fazia *

	1	2	3	4	5	6	7	
Discordo fortemente	○	○	○	○	○	○	○	Concordo fortemente

Source: adapted from Biocca, Harms and Gregg (2001)

Figure G.2: Networked Minds Measure of Social Presence (pt-br) (p2)



Assistência Mútua

Eu não ajudei muito o outro participante *

	1	2	3	4	5	6	7	
Discordo fortemente	○	○	○	○	○	○	○	Concordo fortemente

O outro participante não me ajudou muito *

	1	2	3	4	5	6	7	
Discordo fortemente	○	○	○	○	○	○	○	Concordo fortemente

Eu trabalhei com o outro participante para completar a tarefa *

	1	2	3	4	5	6	7	
Discordo fortemente	○	○	○	○	○	○	○	Concordo fortemente

O outro participante trabalhou comigo para completar a tarefa *

	1	2	3	4	5	6	7	
Discordo fortemente	○	○	○	○	○	○	○	Concordo fortemente

Ações dependentes

Eu não conseguia agir sem o outro participante *

	1	2	3	4	5	6	7	
Discordo fortemente	○	○	○	○	○	○	○	Concordo fortemente

O outro participante não conseguia agir sem mim *

	1	2	3	4	5	6	7	
Discordo fortemente	○	○	○	○	○	○	○	Concordo fortemente

ENVIAR

Source: adapted from Biocca, Harms and Gregg (2001)