GETTING OUT OF FLATLAND

INTERACTION AND INTROSPECTION WITH TANGIBLE AUGMENTED OBJECTS

BY

RENAUD GERVAIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
OF THE DEGREE OF DOCTOR OF PHILOSOPHY
IN THE GRADUATE SCHOOL OF MATHEMATICS AND COMPUTER SCIENCES

université de BORDEAUX

Renaud Gervais: *Getting Out of Flatland*, Interaction and Introspection with Tangible Augmented Objects, © 2015

SUPERVISOR:
M. Martin Hachet, Chargé de Recherche, Inria Bordeaux Sud-Ouest

PRESIDENT:
M. Pierre-Yves Oudeyer, Directeur de Recherche, Inria Bordeaux Sud-Ouest

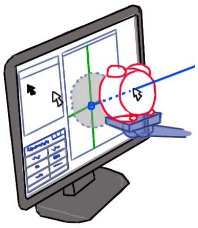
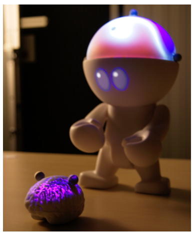
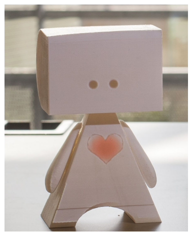
EXAMINERS:
Mme Laurence Nigay, Professeur, Université Joseph Fourier, Grenoble
M. Sriram Subramanian, Professeur, University of Sussex, UK
M. Pierre Dragicevic, Chargé de Recherche, Université Paris Sud

LOCATION:
Bordeaux, France

DEFENDED ON:
December 9th, 2015

Dedicated to the three loves of my life,
Odette, Joline and Dominique,
and to my father, Paul.

ABSTRACT

Most of our waking hours are now spent staring at a screen. While the advances in touch screens have enabled a more expressive interaction space with our devices, by using our fingers to interact with digital content, what we see and manipulate on screen is still being kept away from us, locked behind a glassy surface. The range of capabilities of the human senses is much richer than what screens can currently offer. In order to be sustainable in the future, interaction with the digital world should leverage these human capabilities instead of letting them atrophy. One way to provide richer interaction and visualization modalities is to rely on the physical world itself as a host for digital content. Spatial Augmented Reality provides a technical mean towards this idea, by using projectors to shed digitally controlled light onto real-world objects to augment them and their environment with features and content. This paves the way to a future where everyday objects will be embedded with rich and expressive capabilities, while still being anchored in the real world.

In this thesis, we are interested in two main aspects related to these tangible augmented objects. In a first time, we are raising the question on *how to interact* with digital content when it is hosted on physical objects. As a basis for our investigation, we studied interaction modalities that leverage traditional input and output devices found in a typical desktop environment. Our rationale for this approach is to leverage the experience of users with traditional digital tools – tools which researchers and developers spent decades to make simpler and more efficient to use – while at the same time steering towards a physically enriched interaction space. In a second time, we go beyond the interaction with the digital content of augmented objects and reflect on their potential as a humane medium support. We investigate how these augmented artifacts, combined with physiological computing, can be used to raise our awareness of the processes of our own bodies and minds and, eventually, foster *introspection* activities. This took the form of two different projects where we used tangible avatars to let users explore and customize real-time physiological feedback of their own inner states.

RÉSUMÉ

La plupart des métiers du travail de l'information requièrent maintenant de passer la majeure partie de nos journée devant un écran. Ceci s'ajoute au temps déjà consacré à ce médium dans nos temps libres pour le divertissement et la communication, que ce soit en utilisant des téléphones intelligents, tablettes ou ordinateurs. Alors que les avancées technologiques dans le domaine des écrans tactiles nous ont permis d'interagir avec ces appareils de manière plus expressive, par exemple en utilisant nos doigts pour interagir directement avec le contenu numérique, ce que nous voyons et manipulons sur écran reste "intouchable" ; nos doigts ne pouvant pénétrer la surface de l'écran pour toucher le contenu numérique qui se trouve derrière. Pour que l'interaction avec le monde numérique soit écologique dans le futur, elle doit mettre à profit l'ensemble des différentes capacités de l'humain au lieu de ne se concentrer que sur certaines d'entre elles (comme le toucher et la vision), laissant les autres sens s'atrophier. Une façon de considérer le problème est d'utiliser le monde réel physique comme support pour le monde numérique, permettant d'imaginer un futur où les objets du quotidien auront de riches et expressives fonctions numériques, tout en étant ancrés dans le monde réel. La réalité augmentée spatiale est une modalité permettant d'aller dans cette direction.

Cette thèse s'intéresse principalement à deux aspects en lien avec ces objets tangibles augmentés. Dans un premier temps, nous soulevons la question de *comment interagir* avec du contenu numérique lorsqu'il est supporté par des objets physiques. Comme point de départ de notre investigation, nous avons étudié différentes modalités qui utilisent des dispositifs d'entrée/sortie typiquement retrouvés dans un environnement de bureau. Cette approche est justifiée par le désir d'utiliser au maximum l'expérience que les utilisateurs ont déjà acquise avec leurs outils numériques tout en se dirigeant vers un espace d'interaction comprenant des éléments physiques. Dans un second temps, nous sommes allés au delà du thème de l'interaction avec le contenu numérique pour se questionner sur le potentiel des objets tangibles augmentés comme support pour un médium plus humain. Nous avons investigué comment ces artéfacts augmentés, combinés à différents capteurs physiologiques, pourraient permettre d'améliorer notre conscience des processus internes de notre corps et de notre esprit, pour éventuellement favoriser l'*introspection*. Cette partie a pris la forme de deux projets où un avatar tangible a été proposé pour laisser les utiliseurs explorer et personnaliser le retour d'information sur leurs propres états internes en temps réel.

PUBLICATIONS

The work presented in this thesis has previously appeared in different publications, which are listed below.

- Renaud Gervais, Jérémy Frey, Alexis Gay, Fabien Lotte, and Martin Hachet. TOBE: Tangible Out-of-Body Experience. In *Proceedings of the 10th International Conference on Tangible, Embedded and Embodied Interaction*, TEI '16, Eindhoven, Netherlands, feb 2016. ACM (co-first authorship)

- Renaud Gervais, Joan Sol Roo, and Martin Hachet. Tangible Viewports: Getting Out of Flatland in Desktop Environments. In *Proceedings of the 10th International Conference on Tangible, Embedded and Embodied Interaction*, TEI '16, Eindhoven, Netherlands, feb 2016. ACM

- Joan Sol Roo, Renaud Gervais, and Martin Hachet. Inner Garden: an Augmented Sandbox Designed for Self-Reflection. In *Proceedings of the 10th International Conference on Tangible, Embedded and Embodied Interaction*, TEI '16, Eindhoven, Netherlands, feb 2016. ACM. Work-in-Progress

- Renaud Gervais, Jérémy Frey, and Martin Hachet. Pointing in spatial augmented reality from 2d pointing devices. In *INTERACT*, page 8, 2015

- Jérémy Frey, Renaud Gervais, Stéphanie Fleck, Fabien Lotte, and Martin Hachet. Teegi: tangible eeg interface. In *UIST*, pages 301–308. ACM, 2014 (co-first authorship)

- Renaud Gervais, Jérémy Laviole, Asier Marzo, and Martin Hachet. The Good, the Bad and the Hacked: Creative Coding on Objects. Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction Work-in-Progress, feb 2014. URL `https://hal.inria.fr/hal-01096299`

Drop by drop is the water pot filled.
Likewise, the wise man,
gathering it little by little,
fills himself with good.

— The Buddha

ACKNOWLEDGMENTS

When I finally decided to embark on the journey towards a PhD, it was motivated by the desire to keep growing as an individual and push the boundaries of my own knowledge about the world. This has proved to be a worthwhile pursuit and I admit that this broadened my view of research much more than I expected. I did not only learn about different topics, but also got to meet great and inspiring persons that all contributed to this journey in their own ways.

I would obviously like to first thank my advisor, Martin Hachet, for giving me the incredible opportunity to come to France and work at Inria in the Potioc team. Martin is a very humane person and was very supportive during these three years. He gave me the freedom to explore topics that appealed to me and helped me bring ideas to life. I would also like to thank Jacek Jankowski and Christian Muehl for their guidance and friendship. The many discussions we had about research and life in general really helped me grow both as a person and as a researcher. The years we spent together, along with Iza and Lisa, are full of nice memories I will treasure for years to come. Also, I would like to extend my thanks to Jérémy Laviole and Thomas Hulin for their friendship and support, especially at the start of my PhD. Jérémy is the one that got me into Spatial Augmented Reality along with opening my mind towards art and my many discussions with Thomas on self-improvement has been very inspiring for me. They both made themselves available to discuss a variety of problems with me and helped me sort them out.

Of course, the different projects highlighted in this thesis could not have been possible without the collaboration and help from my friends and colleagues at the Potioc group. First and foremost, I want to thank Jérémy Frey, a person one can only qualify as the ideal collaborator. His enthusiasm, positivity and his seemingly endless reserve of energy has made working with him both motivating and inspiring. I can only hope that our paths may cross again in the future. I would also like to thank Joan Sol Roo with whom I worked in order to create the Tangible Viewports system. His great creativity and intelligence, as well as his technological proficiency made it a joy to work with him. Beyond this, he is also a great individual with whom I enjoyed my everyday interaction and rich philosophical discussions.

Other members of Potioc are also worth mentioning for sharing with me their everyday good mood and friendship. I want to thank Fabien Lotte, Anke Brock, Pascal Guitton, Julia Chatain, Camille Jeunet, Damien Clergeaud, David Furio, Nicholas Katzakis and Asier Marzo.

On a more personal level, I want to thank my family and friends from Quebec for whom a thesis meant seeing each others much less frequently. Every time I end a call with each one of them, I feel like an amazingly rich person. Especially, I want to thank my parents, Odette and Paul, as well as my sister, Joline, for everything. Their love and support have made all of this work possible.

Finally, and certainly not the least, I would like to thank my dear and lovely wife Dominique. Dominique agreed to marry me and join this adventure to France for more than 3 years, knowing very well what doing a thesis would imply. The path was certainly not always easy, but I feel that the past three years were not only filled with opportunities to grow, individually and together, but also with smiles and fond memories. I am blessed to have her in my life and thank her for everything.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

Part I

GETTING OUT OF FLATLAND

1 INTRODUCTION

Most of our waking hours are now spent staring at backlit rectangles of varying sizes. While the use of computational devices has become ubiquitous in our daily lives, flat and rectangular screens are still the de facto channel for consuming and interacting with digital content. While they serve a great purpose at delivering this content to their viewer(s), they fail to seamlessly merge with the environment in which they exist. Instead, the user is required to divert his attention from the real-world before him to focus temporarily or for an extended period of time on this flat rectangle. He consumes the information or complete a task and *then* come back to the real-world context.

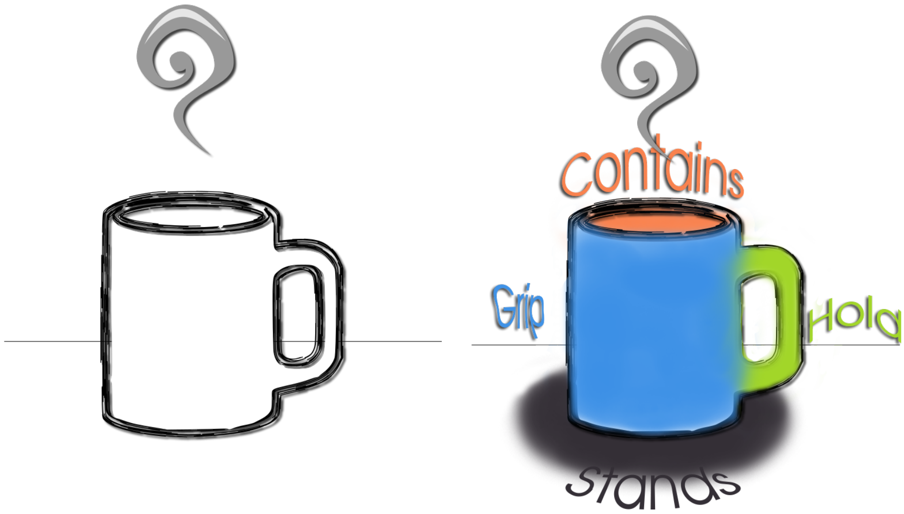
There are multiple problems with this way of handling the meshing of our digital and real lives. For one, our digital lives exist in completely isolated states. We carry them around in our pockets, but one of our scarcest resource, our attention, has always to be either on the digital world or the real one. Additionally, the interaction with the different screens that populate our lives is still very limited. While the advances of touch screen technologies have enabled us to use our fingers to interact with the digital content displayed on screen, we can hardly make the case that it leverages all the rich possibilities of the human hands [165]. In his seminal talk, "A Humane Representation of Thoughts" [167], Bret Victor makes a convincing case for designing interfaces that leverage the vast amount of human senses *as ways of thinking*. He argues that technology right now confines us in thinking only using ways that a computer can easily understand while we have evolved to think *with* a multitude of different senses that involves our whole body. This is limiting. Not only that, it also prevents good ideas to come to fruition because digital matter is right now stiff and “glassy”. You cannot shape an idea with your hands freely; you have to comply to the communication channels which a computer can understand – which most often are different that the ones we would use “naturally”. In order to create a humane medium – a medium enabling and encouraging the different human capabilities and ways of thinking – , focusing on the real world itself seems a promising idea: if we have evolved to think within a real, physical world, leveraging these millions of years of evolution should not only be logical but a sustainable option as well. Indeed, instead of letting our bodies sit at a desk all day to atrophy while our head and fingers do all the work, we should put them to use to instead help us think new ideas. As a way to contribute to this vision, we will focus on physical objects and their use in combination with technology.

1.1 AUGMENTED OBJECTS

In this thesis, we are interested in the use of physical objects to interact with digital information as a way to go *beyond the screens*. We do not argue that screens are devoid of any interest, but when considered as the only mean to convey digital information, they are *limiting*. Different research visions consider the use of computational objects. *Tangible User Interfaces* is a vision where digital information can be handled physically through "Tangible Bits" [65]. *The Internet of Things* is a paradigm in which objects are able to interact with each other and cooperate to achieve common goals [3]. *Organic User Interfaces* (OUI), defined by Holman et al. [59], describe a future where thin and flexible displays as well as touch sensors will wrap everyday objects. Therefore, any part of an object can be a sensor and/or a display. *Augmented Reality* consists in overlaying a live view of the reality with computer-generated graphics to form a coherent and augmented view of the reality [6]. We position ourselves along these disciplines and are interested in physical objects that are *embedded* with computational properties. That is, augmented objects.

Augmented objects have the potential to conciliate the flexibility of the digital world with our expertise of the real world. Note that real-world skills are not innate. However, we get to hone them naturally, simply because our bodies are real-world entities. For example, just by looking at a mug (Figure 1), you can already infer a lot of its properties and also its possible uses. It's hollow and the material it is made of does not seem porous, allowing it to contain matter – matter that potentially would be difficult to hold otherwise, like liquids. The handle seems a good fit for your hand to hold it. Alternatively, the cylindrical shape of the main body could also be grasped easily. The term *affordance* was proposed by Gibson [44] to refer to the actionable properties between the world (object) and an actor (user). Objects not only convey affordances but also conventions, as introduced by Gibson [45] and Norman [114]. Conventions are often cultural and learned. Whereas they do not physically prevent an activity, they prohibit some activities and encourage others [115]. For example, in most western countries, the red color will indicate that an action is forbidden or not recommended as the green color will suggest something that is allowed.

If we consider a future where our daily objects are augmented with digital functions and appearances (Figure 2), it raises questions about the content hosted by these objects. What should be displayed, how can this content be created, by whom? Holman et al. [61] argue that object designers will have a central role in this process and will



(a) A mug. (b) Possible perceived affordances.

Figure 1: Affordances are perceived actionable properties of an object. For example, a mug can be held by its handle, or by its round body that is approximately the size of the hand. It is hollow, so it can be used as a container. Its base is flat so that it can stand on the table.

need new tools more adapted to organic design. Nonetheless, we think that object designers are only one part of the equation. As demonstrated by the rise in popularity of the "maker", "hacker" and "Do-It-Yourself" cultures, users are not only interested in consuming content but also in producing it. Thus, another question that will be of great interest when these objects become widely available is *how to create content*.

In this dissertation, we explore this question using different angles. As a first step, we investigated the *interaction* with augmented content in the context of standard input devices – mouse and keyboard – and desktop environments (Figure 3). The main motivation for this approach is to leverage the use of already existing digital tools that are already known of content creators. For example, programming and visual design are activities well suited for desktop platforms. In the same way that we are interested in anchoring interaction in the real world because users are already very skilled in interacting with its elements, starting our investigation with tools that users already spent years getting proficient with – and which researchers and developers spent decades to make simpler and more efficient to use – seems a promising approach towards a more reality-based approach to computing.

Figure 2: A sketch of an augmented mug. It displays the remaining steeping time for a perfect cup of tea. The overlay shows the quantity and temperature of the liquid inside. The handle will glow green when the tea is ready and that the temperature is appropriate to avoid nasty tongue burns.

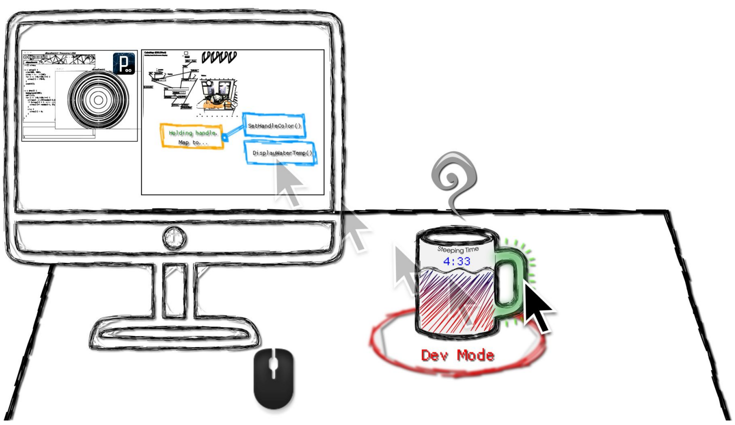


Figure 3: A hypothetical development environment for augmented objects using a standard desktop computer screen and mouse.

1.2 KNOW THYSELF

In a second time, we go beyond the interaction with the digital content of augmented objects and reflect on their potential as a humane medium support. As the form factor of our computational devices gets increasingly personal – smartphones and smartwatches are quickly becoming ubiquitous and are with us all the time – it is important to reflect on the effect they have in our lives. For the most part, they are greatly beneficial as they allow us to stay connected with loved ones, access to a wealth of information, work, play, etc. However, this constant ability to be connected – or inability of being alone [156] – leaves little time for our minds for calmness and contemplation. Email is operating 24h a day while we clearly cannot. While calmness and contemplation might appear to be important only for an overall feeling of well-being – which is very important on its own –, it is also central in fostering creativity and innovation. In his book "Where good ideas come from: The natural history of innovation" [70], Johnson makes the case that great ideas often comes from *slow* hunches that, therefore, require *time*. Richard Hamming, in his seminal talk "You and your research" [49] hints in the same direction when mentioning his habit to set time aside weekly for thinking about "Great Thoughts". It is indeed hard to imagine any Nobel prize laureate reflecting on important problems while answering emails or disturbed by yet another notification from his or her smartphone.

Augmented objects, especially tangible ones, have the potential of creating experiences that both complement the real world and provide an accessible way to interact with its content. Augmented reality has the capability to expose hidden information about the world. More specifically, it can be used to gain more insight about one of the most personal aspect of lives: our own selves – bodies and minds. Indeed, while we are using our bodies and minds every day of our lives, we are most often unaware of their inner working and states. Physiological computing [34] is getting mature enough to not only measure basic physiological signals such as heartbeats and breathing, but also mental states [123].

In the second part of this thesis, we investigated the use of augmented objects as a way to foster *introspection*. We explored ways to allow users to create and investigate representations of their inner selves by using physiological sensors – heartbeats, breathing, electroencephalography.

1.3 CONTEXT

This section aims at giving an overview of the research areas onto which this thesis builds. It mainly relates to Ubiquitous Computing (Ubicomp), Physiological and Affective Computing, Tangible User Interfaces (TUI), Organic User Interfaces (OUI) and Augmented Reality (AR) – more specifically *spatial* AR. The common thread unifying these fields are their goal of weaving the real and the digital worlds together, in different ways.

1.3.1 *Ubiquitous Computing and Calm Technology*

In 1993, in a special issue of the Communications of the ACM, Wellner et al. [181] suggested going towards *computer-augmented environments* in opposition to virtual reality (VR). VR seeks to replace the world in which we evolve daily with a simulated one. Instead, computer-augmented environments would merge both the real world and computational devices into a complete, real-world oriented experience. Mark Weiser, widely recognized as the father of Ubiquitous Computing, was well part of this movement of computers disappearing into the environment [176]. In the same issue of the Communication of the ACM, he published what would be the itinerary of the ubiquitous computing for the following decade [177]. He would later refine this vision by proposing the concept of "calm technology" [178]. The main idea of calm technology is that technology should disappear in the surroundings of users, intervening only when required before "getting out of the way" again. Weiser envisioned a type of technology that would support a state of calmness for people, enabling them to navigate through the turmoils of daily life smoothly. This view has been criticized by Rogers [133] for depicting the users as passive actors while the technology decided when was the best time for it to intervene. Instead, she proposes an alternative view where the users are proactive toward technology: being empowered by technology in their own terms. Weiser's vision also depicts an environment that would inform while in the background by displaying information relying on the periphery of the user's attention as proposed by Buxton [23] – without having to focus on it deliberately. Ishii et al. [66] used this concept in conjunction with architectural space to design a room where information was represented as subtle visual and auditory cues. These work are interesting because they do not claim our limited attention in order to work well. In the same line of thought, Hallnäs and Redström [48] proposed the idea of a technology that is purposely *slow* and is designed to foster *reflection*.

1.3.2 *Physiological and Affective Computing*

Technology is built by and for human individuals. However, too often, the interaction between the technology and said individuals take the form of psychopathic monologues – a serie of commands and feedbacks disregarding any distress, frustration or pleasure the users express. This is because of the inability of machines to understand

and express human emotions. This realization motivated Rosalind Picard to develop the concept of *Affective Computing* [123] along with a whole research agenda. Picard defines this concept as "computing that relates to, arises from, or deliberately influences emotions". Among the many potential technical ways to understand human emotions is *Physiological Computing*. Fairclough et al. [34] define it as a type of computing which uses real-time psychophysiology to represent the state of the user – e.g. cognitions, motivation, emotion. It then uses this knowledge of the user's internal state to adapt the interactive system in real-time. Physiological computing relies on different types of bodily signals, such as heart rate, breathing or electrodermal activity that can be detected in different ways. It is also possible to measure brain activity using varying methods; one of the most popular one consists in measuring the electrical current at the surface of the scalp – electroencephalography (EEG). Measuring brain activity as a communication and control channel for interactive systems is a research area on its own: Brain-Computer Interfaces (BCI) [188, 187].

1.3.3 *Tangible User Interfaces*

Tangible User Interfaces evolved from the trend initiated by the discussion of the "Back to the Real World" issue [181]. While all the pieces were there, it took a few years before it developed as an interaction style. Often cited as one of the first incarnations of TUI is the marble answering machine [28] where marbles embodied voice messages received while the recipient was away from home. Putting the marble in a dedicated groove on the device would play back the message and propose to call back the original caller. Fitzmaurice et al. [36] laid the foundation of the field with its *Graspable Interface*. A few years later, Ishii and his collaborators proposed a defined vision as *Tangible Bits* [65]. The transition from *graspable* to *tangible* was deliberate. Graspable UI emphasized the ability to see and manipulate digital data from physical handles. Tangible bits also encompassed architectural and ambient feedback, emphasizing the use of the whole real world as an interface. During the next decade, Ishii and his students at the MIT Tangible Media Group pursued this vision and advanced the field significantly [64]. More recently, Ishii and his colleagues have updated their research agenda to what they call "Radical Atoms" [67]. Radical atoms is a vision that proposes a shift from tangible interaction to *material* interaction. It hypothesizes a new physical material which physical properties – rigidity, weight, volume, etc. – can be digitally controlled. This vision is definitely the driving force behind Victor's agenda for a dynamic medium of thoughts [167] mentioned earlier.

1.3.4 *Organic User Interfaces*

Organic User Interfaces (OUI) are definitely related to the Radical Atoms vision [67]. Where TUIs have been mostly about interacting with a collection of rigid objects, OUIs emerged from an attempt to move towards more organic and malleable materials. This field of research is mainly driven by advances in flexible input and output technologies. This vision refers to interfaces using non-planar displays that can change shape via continuous physical inputs, either actively or passively [26, 59]. OUIs differ from TUIs in that their surface is always coated with a high-resolution display [164]. OUIs are defined by three main themes:

1. Input equals output: the display is the input device;
2. Function equals form: the display can take any shape;
3. Form follows flow: the displays can change their shape.[59]

In this thesis, we are especially inspired by the second aspect of this definition, since we want to use real-world objects not as proxy or handles but as real objects which form convey a meaning. A few years after giving a definition for these types of interface, Holman et al. [61] published a set of guidelines for the design of such interfaces. Namely, they present the notion of *hypercontext* which is very much linked to the "functions equals form" part of the OUIs definition. Hypercontext is the idea that the interactive behavior of some types of organic designs should express only a small amount of essential actions, which are strongly linked to their form factor. As an example, they consider a credit card which is wrapped by thin interactive displays. The back of the card would exhibit a map indicating the nearest ATM machine or the current account balance because it relates strongly to the card's initial function. It would avoid nesting a web browser or a news reader. In the same article, they also highlight the difference between the OUI and Radical Atoms visions. They argue that Radical Atoms consider shape *as a volume* composed of reprogrammable particles that are used to fluidly represent information while OUIs consider actuation in the context of an object's overall shape – one that is wrapped with an interactive display.

1.3.5 *Spatial Augmented Reality*

Augmented Reality (AR) consists in overlaying computer generated information to our real world experience. The concept has first been described by Sutherland in his description of a hypothetical "Ultimate Display" [148]. This display would be a room that have the power to control matter itself, resulting in a complete weaving of the real and the virtual. Traditionally, the visual information overlay

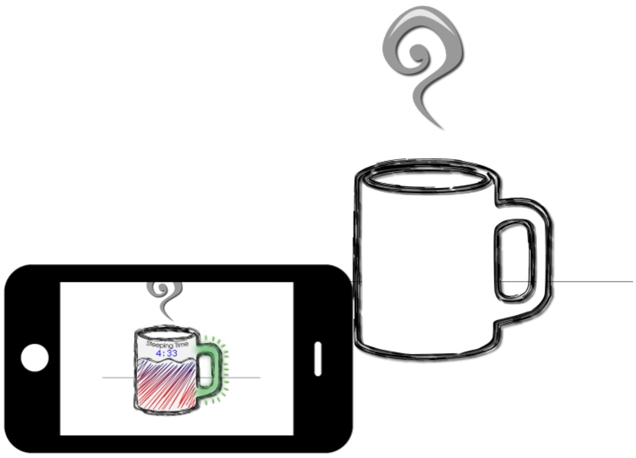
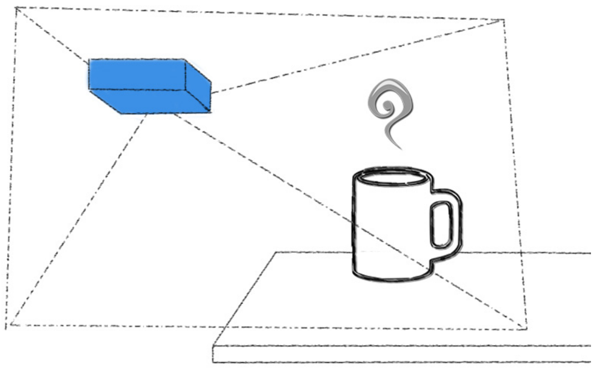


Figure 4: AR is often achieved using a video see-through technique consisting in overlaying computer generated graphics over a live video feed of the real-world. Here, a smartphone is used to produce an augmented mug view.

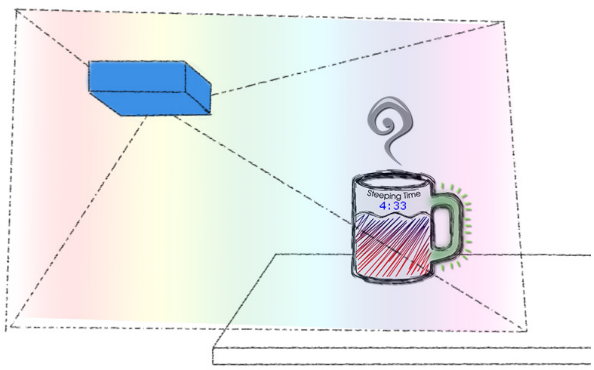
has been conveyed using see-through or video see-through AR. See-through AR consists in having a semi-transparent display which allows to both display the virtual elements and let the user see through it to observe the real world scene simultaneously. Video see-through consists in augmenting a live video feed that is then viewed either using a head mounted display or a mobile device equipped with a camera – e.g. a smartphone or tablet (Figure 4).

An alternative to these standard techniques appeared at the end of the 90s. In 1998, Raskar and his colleagues proposed a new paradigm called *Spatial Augmented Reality* (SAR) [128] which used projectors or displays to create images that are integrated in the user's *environment* (Figure 5). This paradigm is based on the vision laid out by their previous "Office of the Future" publication [127]. A few years later, Raskar et al. [129] applied the same set of techniques to change the appearance of physical objects. The main advantage of SAR lies in the fact that it does not require the user to wear any special equipment – objects and spaces are augmented in place. This also eases a shared experience between multiple users.

All the work presented in this dissertation uses SAR to a certain extent. It does not mean that SAR is necessarily the best choice for creating augmented objects as final products. However, it represents a flexible way to create and interact with such objects. For this reason, Chapter 2 is reserved for going in further details about this technological choice's inner working and background.



(a) A normal mug with the projector turned off.



(b) A projector augmenting a normal mug with digital content.

Figure 5: SAR is achieved using projectors or displays to create images that are integrated in the user's environment.

1.4 CONTRIBUTIONS

In this thesis, we explore methods and systems to interact with tangible augmented objects. Additionally, we investigate ways to leverage these augmented objects in combination with physiological computing to foster introspection. More specifically, our contributions are the following:

1. The evaluation of the use of 2D pointing devices – mouse and graphics tablet – in a pointing task in a SAR context compared to a screen condition.
2. The design, implementation and evaluation of a system enabling the interaction between a typical desktop computer environment – traditional screens, mouse and keyboard – with tangible augmented objects, considering an object design scenario as a main thread.
3. The design, implementation and evaluation of a tangible interface enabling users to visualize and interact with their real-time electroencephalography (EEG) readings in order to explore different brain processes.
4. The design, implementation and evaluation of a toolkit enabling users to create tangible augmented avatars providing feedback from real-time low- and high-level physiological readings.

1.5 STRUCTURE OF THE THESIS

This thesis is divided into four parts. Part i lays out the overall motivation and context of this research (Chapter 1) and provides details on the topic of Spatial Augmented Reality (Chapter 2) since it is the technological thread unifying the different projects. Note that we purposefully avoid a dedicated chapter for related work as we instead present them as appropriate alongside the work of each chapter.

Part ii focuses on the *interaction* with the digital content of tangible augmented objects. Chapter 3 present a study evaluating the performance of a pointing task on spatially augmented objects using standard 2D input devices such as mice and graphic tablets. We were interested to see if it was possible to use the same pointing technique used in standard computing with a real world scene. Chapter 4 introduces the Tangible Viewports system which builds onto the pointing task study presented in Chapter 3. A Tangible Viewport is an on-screen window that allows interaction between a desktop computer application and an augmented object placed in front of it, using standard input devices. Effectively, it recreates the interactive behavior of a standard viewport rendering a 3D scene with the main exception that the 3D scene is composed of a real, tangible object. The goal with this system is to leverage the power and flexibility of standard computer applications for content creation while enriching the content creators with physicality in their workflow. We also present an exploration of the design space enabled by this way of interacting with digital content hosted on physical objects.

In Part iii, we present two different systems using augmented objects as a way to foster *introspection* – getting to know more about oneself and one's internal physiological and mental processes. Chapter 5 present a friendly little character named Teegi – Tangible Electroencephalography (EEG) Interface – a tangible interface to learn more about the internal working of the brain. EEG is a technology that measures the electric current on the surface of the scalp. It allows to infer the electrical activity of the brain and is commonly used in Brain-Computer Interfaces (BCI). Teegi itself is an anthropomorphic character onto which a user's live EEG reading is displayed. The user can manipulate Teegi to explore the signals and change the different filters applied to the EEG readings using small tangible characters – mini Teegis. Expanding upon the work with Teegi, we created a broader toolkit to allow anyone to actively and physically build a tangible representation of one's inner state that is detailed in Chapter 6. This toolkit is named TOBE – to be pronounced [tobi] – a Tangible Out-of-Body Experience. This work encompasses the whole workflow: creation of a tangible support using 3D printing, building the physiological sensors (e.g. heart rate, breathing, EEG) using open hardware, the signal processing pipelines and the creation of custom visual feedback.

The work presented in Part iii has been realized in close collaboration with Jérémy Frey[1]. It is important to emphasize and recognize this collaboration since the projects presented in Chapter 5 and 6 could not have been possible without the skills and knowledge of the both of us. Jeremy brought to the table a deep knowledge of physiological computing, electronics and signal processing. My personal contributions to these projects were focused on the tangible, interactive and visual aspects of the systems.

Finally, Part iv concludes this thesis by offering a high-level view of the contributions as well as discussing potential future works. Figure 6 provides a visual overview of this thesis plan.

1 http://jfrey.info

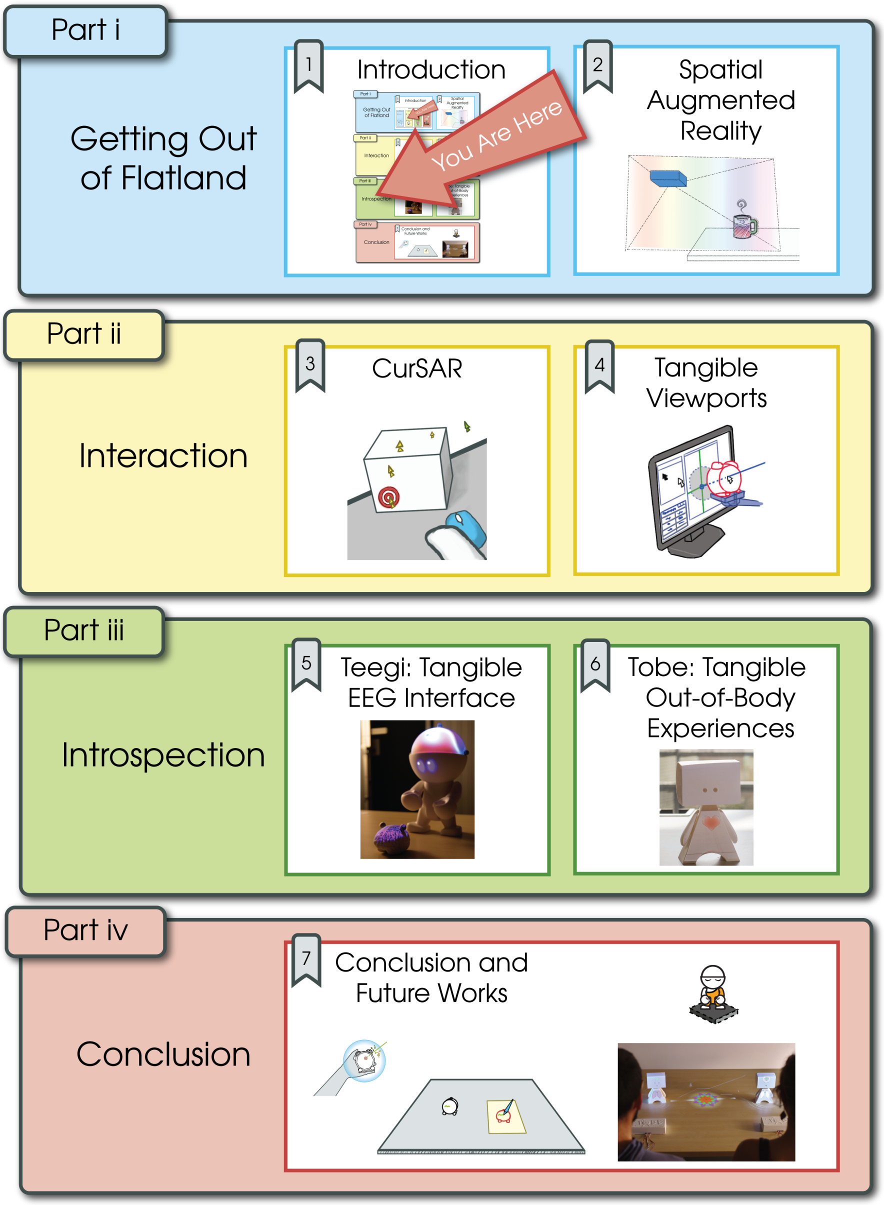


Figure 6: Plan of the thesis

2 SPATIAL AUGMENTED REALITY AS A MIXING MEDIUM

This chapter provides an overview of Augmented Reality (AR) and, more specifically, Spatial Augmented Reality (SAR). Since the work presented in all following chapters uses SAR, this chapter aims at pooling the related work and implementation details in a single place in order to avoid useless repetitions. The goal is also to explain enough background to provide a global understanding of SAR principles and some details on the implementation used throughout the thesis. While we will cover some basic notions of SAR, an interested reader can get a more thorough understanding of the different principles of this AR paradigm in Bimber and Raskar's book on the topic [19].

2.1 BACKGROUND

This section covers the main concepts and definitions related to AR and SAR. AR is first defined and a brief overview of its history is given. Then we will focus on one of its subsets: SAR.

2.1.1 *Augmented Reality*

As mentioned in Chapter 1, augmented reality consists in overlaying computer generated information on a real-world experience. The overarching goal of AR is to create a seamless experience, merging both real and digital information. Sutherland [148], as soon as 1965, proposed the idea of an "Ultimate Display": a room within which the computer could control the existence of matter. A few years later, he was also the first to build an AR system, taking the form of a head-mounted display (HMD) [149]. However, using a HMD is only one way to visually augment the real world. Azuma later provided a definition of AR that is flexible enough so not to be technology dependent [6]. An AR system would then have the following properties:

1. Combines real and virtual
2. Is interactive in real-time
3. Is registered in three dimensions

AR is neither purely a real-world experience nor is it pure virtual reality. Milgram and Kishino [104] proposed a taxonomy for mixed reality display technologies taking the form of a continuum between

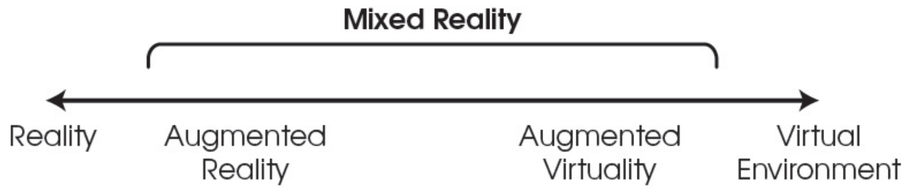


Figure 7: Milgram's virtuality continuum as described in [104].

reality and virtual reality as shown in Figure 7. At the left extremity is the unaltered reality. On the other side lies virtual reality, which replaces reality completely relying on different sensory channels – usually visual and auditory. Augmented reality is located on the left-hand spectrum of mixed reality. Indeed, AR is centered on real-world elements onto which digital information is added. Note that Milgram, in his virtuality continuum, did not explore the different display devices for AR.

This proximity to reality is very well illustrated by the title of the special issue of the *Communication of ACM* in 1993 – "Computer Augmented Environments: Back to the Real World" – which contributed in establishing the AR research agenda [181]. In 1998, Mackay [95] reflected back on this issue and described the three basic approaches to augmenting real-world objects:

AUGMENT THE USER: The user has either to wear or carry a device which is used to see the augmented information. This usually takes the form of a HMD or a mobile device, using see-through or video see-through approaches.

AUGMENT THE PHYSICAL OBJECT: Physically modify the objects themselves with different input/output components. For example, an object's surface could be covered by a thin film display, such as what is described in the Organic User Interfaces vision (see Section 1.3.4). This hardware modification with computational devices is also reminiscent of the Internet of Things trend [3].

AUGMENT THE ENVIRONMENT: Independent devices are installed in the environment which collect information from their surroundings and display information onto objects. They also handle the users interaction. This approach leaves both the users and the physical objects unaffected. This approach is the one leveraged by SAR which will be explored in more depth in the rest of this chapter.

It is important to note that while the vast majority of the AR research community focused on augmenting the visual sensory channel, it is not restricted to it in any way. For example, Dobler et al. [32] used AR to enable users to see and place sound sources in space. Another example is the imaginary reality game of Baudisch and his colleagues [12]. In this game, two teams played a basketball game in which there is no physical ball to manipulate. Instead, there is an imaginary ball that exists solely via a computer analyzing the playing area in real-time. The location of the ball has to be inferred by the players using sound cues given by the system managing the game.

2.1.2 *Spatial Augmented Reality*

SAR is a subset of AR that consists in displaying the augmented content *in* the environment using projectors or displays. The first steps in this direction have been highlighted in the "Office of the Future" vision [127]. Figure 8 provides a good illustration of the authors' vision. It describes ideas and techniques for creating immersive in-situ displays, building upon the notions of the CAVE system, in a normal office environment. By using computer vision to compute the depth of every pixel of a camera, they could correct the projected images according to the topology of the projection support. Their dream was to have a room in which the light of *every millimeter* could be controlled at *every millisecond*.

Based on the vision of the Office of the Future, Raskar and his colleagues formalized these techniques into a new paradigm to achieve AR, which they named *Spatially* AR [128] – nowadays, it is more often referred to as *Spatial* Augmented Reality or projection mapping in artistic communities. They put emphasis on augmenting the *environment* instead of the users' field of view. By tracking a user's head position, they proposed a set of techniques to project on irregular surfaces in the environment so as to generate perceptual illusions for the user. When updating the projection in real-time, it is possible to make it appear as if virtual objects are registered in 3D to physical objects. This process is often used in arts to create surreal and surprising illusions that only works for a specific viewpoint. It is usually referred to as *anamorphic illusions*. A basic example of the principle is shown in Figure 9. Other works using SAR to produce anamorphic illusions include [88] which used a mobile robot mounted with a projector. More recently, Benko et al. [16] proposed a dyadic SAR system: a room for two users where each person, facing each other, have a view on common virtual elements. Interestingly, the system also uses the users themselves as projection surfaces for the other user's view (see Figure 10).

Projection of digital information in the environment has been done before. Wellner [180] presented his DigitalDesk in 1993 which projected digital documents on physical paper laid out on a desk and with which it was possible to interact with a pen. This work was essentially closing the loop of the Desktop metaphor used ubiquitously in computing since the 1970s, by displaying back the digital desktop on a real desk. Mackay [95] also worked with augmented paper using projection for engineering drawings, video edition and flight control.

(a) A normal office space where part of the walls and table can be used as spatial displays.

(b) When the displays are active, they can be for example used to create a virtual shared office.

Figure 8: Conceptual sketch of the Office of the future, as presented in [127]. Image courtesy of Raskar et al.

(a) (b)

Figure 9: A simple example of anamorphic illusion. *(a)* When the viewpoint of the camera is exactly at the right spot, the image of the augmented mug appears as intended. *(b)* The actual image printed on the sheet of paper used in order to produce the image on the left.

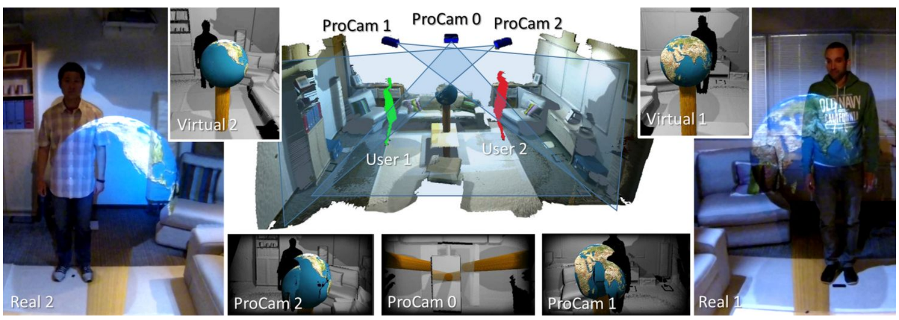
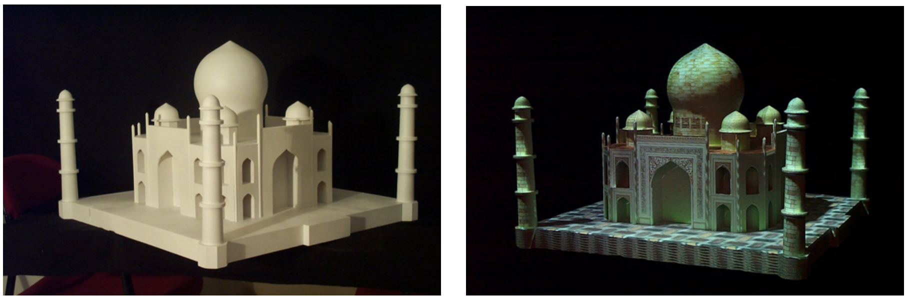


Figure 10: The Mano-a-Mano system using a dyadic projection system to create anamorphic illusions in a room for two users. Image courtesy of Benko et al. [16].



(a) (b)

Figure 11: Using a projected light to change the appearance of a physical object. Image courtesy of Raskar et al. [129].

While not creating augmented objects per se, Underkoffler and Ishii used a combination of tangibles and projection in the environment to create simulations. They presented an optical prototyping tool where various optical elements were controlled by physical objects while the a simulation of the light behavior was projected on a worktable [158]. In *Urp* [159], they re-used the same underlying technology – which they named *I/O Bulb* for its capacity to use light as input and output – to create a urban planning tool which projected simulated shadows of physical wooden structures laid out on a table. Other types of simulations, such as fluid flow, was possible.

SAR can also be used to change the appearance of objects by mapping different textures to its surface (Figure 11). This idea was introduced under the term "Shader Lamps" by Raskar and his colleagues in 2001 [129]. The term "shader" was used to highlight the fact that this technique could be used to create the illusion of different materials and simulate artificial illuminations on the augmented objects. One of the advantages of this technique over anamorphic illusion is that it does not require the users' head position[1]. Therefore, it allows multiple users to see the same augmentation simultaneously. Bandyopadhyay et al. [9] made those shader lamps dynamic, using them in conjunction with a tracked pen to allow users to digitally paint a tracked physical object.

2.2 HOW DOES IT WORK?

In this section, we are taking a closer look at how SAR is actually achieved in practice. While most of the concepts presented here can be generalized from a standard AR pipeline, we will focus on its spatial specialization. As discussed in the previous section, there are mainly two different types of augmentation that can be done with SAR: texture mapping – Shader Lamps – and anamorphic illusions. The creation of augmented objects relies more heavily on texture map-

1 Note that simulating artificial lighting still requires the head position of the viewer.

ping. It is the main technique used in Chapters 3, 4, 5 and 6. However, since Chapters 3 and 4 also slightly rely on anamorphic illusions, they will also be briefly covered in this section.

2.2.1 *Texture Mapping*

In order to create an augmented scene, a virtual counterpart of every real-world component that is used to create the augmented experience is required. Indeed, we want to create the illusion that augmented graphics displayed on the physical object's surface are actually *part of* the object. This first require a virtual version of the object to be augmented. All the digital operations and animations will be created and handled in the virtual world using this virtual object. Then, we model the projection cone of our projectors in the real-world. This consists in characterizing the behaviors of the pixels in space. A virtual camera is then created based on the projector's parameters. We finally reproject this virtual camera view on the real world environment using the projector.

This pipeline and its virtual counterpart are respectively summarized in Figure 12 and 13. It can be broken down in four main components – labeled from 1 to 4 in Figure 12 – which will be described in further details in the following subsections:

1. Geometry of the augmented object.
2. Position of the augmented object in the world: tracking.
3. Position of the projector(s) in the world: extrinsic calibration.
4. How the pixels of the projector are traveling through physical space: intrinsic calibration.

2.2.1.1 *Geometry of the Object*

The first step in order to create an augmented object is to have a 3D model corresponding to the physical object to be augmented. It is possible to approach the problem from two different angles:

- From real to virtual
- From virtual to real

FROM REAL TO VIRTUAL This approach consists in starting from a real object and then creating a virtual version of it. One of the most straightforward way to achieve this is by simply measuring – using, for example, a measuring tape and a protractor – different key points of the object. They are then processed in a 3D Computer Assisted Design (CAD) software to form a more or less complete geometric

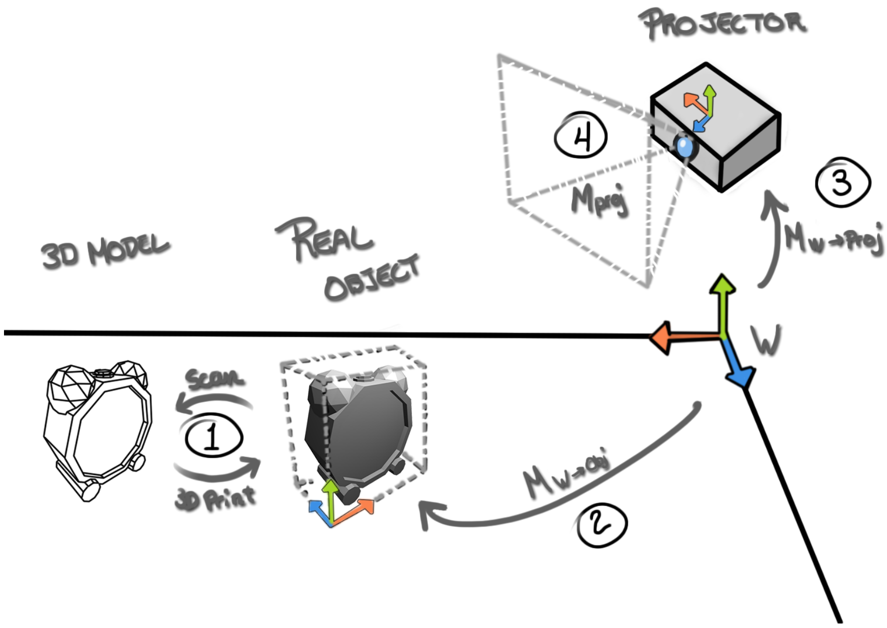


Figure 12: A summary of a SAR pipeline. *W* represents the world's coordinate system's origin. *(1)* The exact geometry of the augmented object must be known, either by using 3D scanning of an existing object or by creating a 3D model and building it using, for example, 3D printing. *(2)* The position and orientation of the object in world space ($M_{W \to Obj}$) must be known via a tracking solution or manual measurements. *(3)* The position and orientation of the projector in world space ($M_{W \to Proj}$) must also be known. *(4)* The intrinsic parameters of the projector (M_{Proj}) is required to know how the pixels are being distributed in the real-world environment.

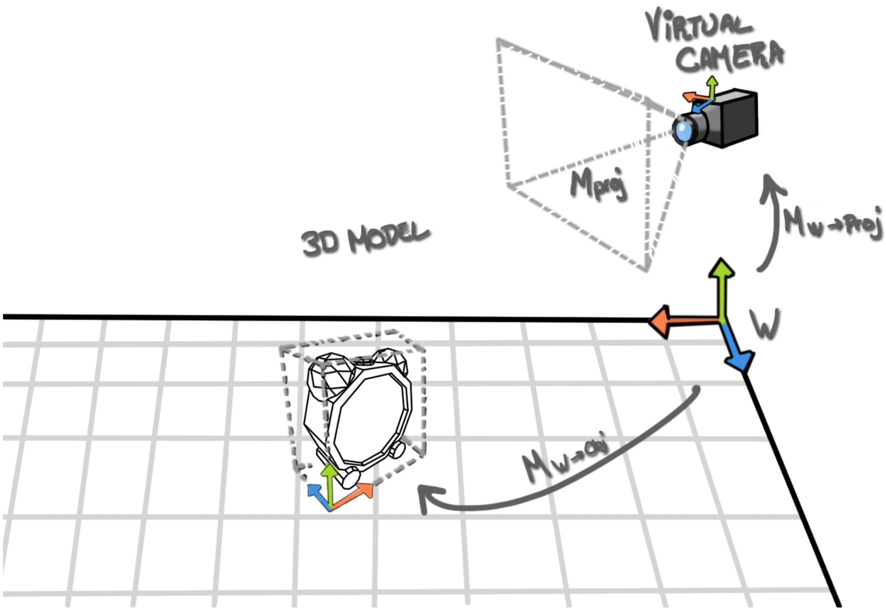


Figure 13: The virtual counterpart of the SAR pipeline shown in Figure 12. The 3D model of the real object is moved to the right location in virtual space $M_{W \to Obj}$ as fed by the tracking solution. Then, a virtual camera is created using the parameters returned from the extrinsic ($M_{W \to Proj}$) and intrinsic (M_{Proj}) calibration of the projector. Finally, the image produced by the virtual camera is sent to the projector to create the augmented scene.

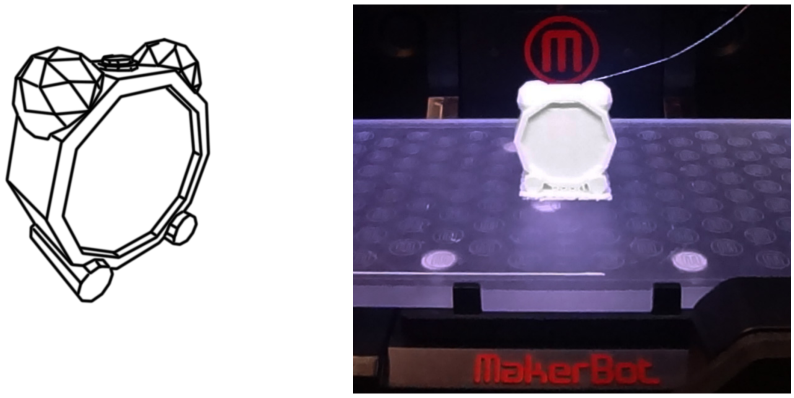
shape. This method works fine if the measurements are made carefully and if the object is relatively simple. However, as the geometry of the object increases, this method becomes painstakingly slow and error prone.

An alternative way to achieve this is to rely on an automatic process that will generate the 3D points for us by scanning the object. Scanning can be conducted in many different ways including using cameras and lasers – such as LiDAR systems. Recent advances in computer vision techniques and the democratization of depth cameras such as the Microsoft Kinect™ has enabled 3D scanning at very low costs [111]. However, the Kinect's sensors are still relatively low resolution and do not operate well below a certain distance – 0.5 meters for the Kinect v2. Consequently, relatively small objects do not produce high-fidelity 3D models. While it might be sufficient for certain types of augmentation, anything related to design and precision applications will suffer greatly from a lack of precision in the 3D model. Scanning can also be achieved using structured light patterns captured by a camera [108]. This method will produce a 3D point cloud which will then need to be meshed.

FROM VIRTUAL TO REAL This method is typically used in industrial settings where object fabrication starts with a digital design. The object is then physically built either manually – for example with clay during the design stages – or with machines such as CNC mills. However, recent advances in digital fabrication technologies, such as 3D printing and laser cutting, has widened considerably the accessibility object making. One of the main advantages of digital fabrication is that it ensures that the physical object corresponds to the 3D model (Figure 14). Moreover, 3D printing also has the benefit of providing some control over the material that is used for the object creation. Indeed, in SAR settings, the properties of the material receiving the projector's light is crucial. That is, a diffuse white material is often best. Moreover, advances in 3D printing technologies allow for more fine grained control over the material being used at very specific locations in the object. For example, translucent glass has been used in combination with 3D printing [78]. More recently, it has even been used as a way to create interactive components by printing embedded optical components [183] and curved displays [21].

2.2.1.2 *Tracking*

Tracking is the process by which the position and orientation – which we will call *pose* from now on – of the physical object to be augmented is retrieved in world space. This corresponds to step 2 in Figure 12. Once the exact geometry of the object is known, we must replicate its real-world pose in the virtual environment. The requirements in precision and performance depend on the application at hand.



(a) (b)

Figure 14: 3D printing allows to create a replica of a digital model.

The precision[2] is typically expressed in meters and degrees although it is often found expressed in pixels when the tracking solution relies on computer vision. Having a precise tracking system – and a good projector calibration (see Section 2.2.1.3) – is important to minimize augmentation artifacts such as bleeding or misregistrations between the augmented content and the physical object (Figure 15). Performance is often evaluated in milliseconds and is indicative of the system's capacity to sustain real-time updates. That is, when the physical object is moved, how much latency is created by the system having to compute the new pose of the object. Naturally, high tracking performance is required whenever the physical object is to be manipulated and moved around a lot. Moreover, delays between a change in the physical object's pose and the augmentation will break the illusion of a coherent world, which we want to avoid as much as possible. On the opposite, some setups require very little manipulation and can even fare well without any dynamic tracking solution, for example if the object never needs to be manipulated physically[3]. Tracking is one the most challenging aspect of the AR pipeline and has been the main focus of the ISMAR community for a long time [33].

There are multiple ways to tackle the tracking problem in a variety of contexts: magnetic sensors, microelectromechanical systems[4], global navigation satellite systems[5], sonar and computer vision, to name only a few. It is however not our goal to go in depth on this topic since it is not the main interest of this dissertation. We will instead only discuss some of the techniques relying on computer vision

2 Note that when talking about precision, we mean it in a way that encompasses both *precision* and *accuracy*. The difference between these two concepts is highlighted in Section A.1.
3 The study reported in Chapter 3 did not require any tracking.
4 MEMS
5 GNSS. GPS is an instance of a GNSS.

(a) Bleeding

(b) Misregistration

Figure 15: Typical artifacts encountered in SAR. *(a)* Bleeding of some of the projected pixels outside of the object's surfaces. Here, bleeding can be seen as a white silhouette on the table. *(b)* Misregistration artifacts where part of the objects are clearly misaligned. The projection on the feet of the clock are offset from their intended position.

since they are the most commonly found for small object tracking in an AR context. The main tracking technology used in the different projects highlighted in this thesis is based on computer vision.

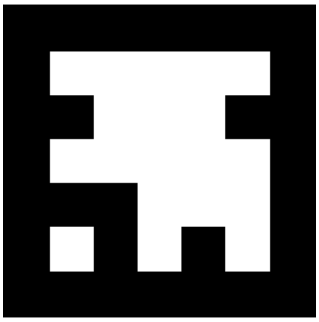
It is important to note that the choice of using SAR imposes certain constraints on the chosen tracking solution. This is because the object being tracked is also the *display*. Therefore, any components attached to the object for tracking purposes should avoid occluding the surface onto which the augmentation will be displayed. Moreover, since we are projecting dynamic content onto the object, any method relying on a static object texture should be avoided.

FIDUCIAL MARKERS The use of markers detected by a camera has become the "visual trademark" of AR. They are comprised of patterns that are easy to detect and identify for computer vision algorithms. Moreover, their square shape, which is easily distinguishable, is used to compute their pose in 6 degrees of freedom (DoF) – position and rotation – in the camera's coordinate system. Enabling libraries in this area are ARToolkit [75] and ARToolkitPlus [172]. An example of marker for each of these libraries is shown in Figure 16. Other libraries – e.g. reacTIVision [73] and BullsEyes [79] – are more specifically designed for tabletop use and are limited to 3 DoF (x, y, θ).

Using fiducial markers requires having a calibrated camera[6]. This process essentially consists of transforming pixel measurements in

6 Camera calibration is a process very similar, yet simpler, to projector calibration (Section 2.2.1.3).

(a) ARToolkit marker [75].



(b) ARToolkitPlus marker [172].

Figure 16: Example of fiducial markers commonly used.

the real-world coordinate system – i.e. meters. This tracking solution has the obvious advantages of being cheap, accessible and can achieve good and robust precision. The markers are required to be planar, therefore it is important that they are not folded or distorted in any way. They also require to be fully visible in the camera's frustum in order to be detected. To overcome this limitation, it is possible to combine multiple markers together that acts like a single tracked target. Their form factor makes them difficult to affix on complex 3D objects. Another drawback is their visual clutter. Indeed, their appearance and form factor can be disturbing for users when manipulating a physical object. Some research projects have successfully reduced the markers' visibility to the naked eye by relying on the infrared spectrum [94, 184].

NATURAL FEATURE TRACKING Image features are characteristic points processed in certain ways which make it possible to detect and match them together, using computer vision algorithms. Commonly used are SIFT [93], SURF [13] or ORB [135]. Natural feature tracking can be used in a variety of ways. For instance, it can be used following the same principles of fiducial marker based pose estimation presented in the previous section. It mainly consists in replacing the very obvious markers, such as the ones in Figure 16, by more natural images such as a logo, a photograph or the natural texture of an object. Professional solutions, such as the Vuforia® framework[7], propose AR solutions based on this technology. Alternatively, natural feature tracking is often used in scenarios where you have to track a camera's pose which is evolving in an environment to be augmented using video see-through techniques. Approaches such as Simultaneous Localization and Mapping (SLAM) [91] can leverage the use of these images features, even when using a single camera [30].

Natural feature tracking is difficult to use in SAR contexts for obvious reasons: by augmenting the object using projected light, we alter

7 http://www.qualcomm.com/products/vuforia

its appearance. Hence, it disturbs the tracking since it is based on the object's texture in order to provide stable pose estimation.

DEPTH-BASED TRACKING This category is relatively broad as it encompasses a lot of different techniques. We will only discuss the ones that have been considered for our work with SAR. It consists in using 3D – or 2.5D – data to track an object in space. As mentioned above, the democratization of depth sensors such as the Microsoft Kinect™ or the Asus Xtion™, has made depth sensing a realistic approach with off-the-shelf components. Generating depth information from 2D sensors such as cameras can also be achieved by using multiple sensors together. First, stereo pairs – two devices that have partially the same view on an environment and that are calibrated together – need to be created. Stereo calibration is done in order to determine the relative positions of the two cameras' frustums. Then, any matching point located in each camera's field of view can be triangulated to determine its depth.

One approach to track object from depth cameras is to compare the 3D model of the object to the point cloud generated by the camera(s). For example, by generating a template point cloud from the 3D model of the object, it could be possible to use point cloud registration techniques, such as the Iterative Closest Point algorithm [18], to register it to the point cloud generated from the camera. While these methods can work in real-time in some cases, they are computationally expensive and the tracking precision is highly dependent on the quality of the camera's sensor. However, higher resolution cameras obviously require more processing power.

Another approach in this category is the use of motion capture systems. These includes professional solutions used in movie studios such as OptiTrack®[8] or Vicon®[9] systems. They work on the same basic principles as the depth sensing cameras – using multiple infrared (IR) cameras to create a tracking volume – but are highly optimized for speed and stability. By tracking only small reflective markers affixed to the physical object, which are illuminated with IR light, these systems can reach refresh rates of 120 frames per second and millimetric precisions. By comparison, most displays refresh between 50-75 per second since most persons have a perceptual time resolution of 50-Hz [174]. Figure 17 shows an example of an object tracked using multiple reflective balls attached to its base. The markers are tracked as a group and therefore tracking is robust to partial occlusions when enough markers and/or cameras are used.

8 http://www.optitrack.com/
9 http://www.vicon.com/

Figure 17: Multiple reflective balls are attached to an object's base. IR light is shed on them and tracked by multiple IR cameras.

2.2.1.3 *Projector Calibration*

Calibration is the process of characterizing the behavior of the light in relation to the optical device we are using. Essentially, a projector can be seen as an inverse camera: it emits light instead of capturing it. Nonetheless, the optical components are more or less the same. However, calibrating a projector is slightly more difficult than calibrating a camera since it has no mean to "see" the real world. In an AR context, our overarching goal is to be able to determine exactly what the projector is "seeing" so that we can create a corresponding virtual camera to correctly display the augmented content. We will therefore have to tackle the problem in an indirect way compared to a camera calibration process.

A projector or camera can typically be characterized by two components: *extrinsic* parameters and *intrinsic* parameters. They respectively refer to steps 3 and 4 of Figure 12. A simplified explanation of these parameters is presented below. An interested reader can refer to Appendix A.2 for more details on the pinhole camera model used here.

EXTRINSICS The extrinsic parameters of a projector simply refer to the position and orientation of the imaging component, in the chosen world coordinate system. This can be modeled as a matrix allowing to convert a point's position, expressed in world coordinates, in the projector's local coordinate system.

INTRINSICS The intrinsic parameters refer to the projection cone of the optical system as well as the modeling of deformation in the image caused by the imperfection of the optics. This can be modeled as a matrix enabling to convert the position of a point, expressed in the projector's coordinate system in meters, in a position on the imaging plane expressed in pixels.

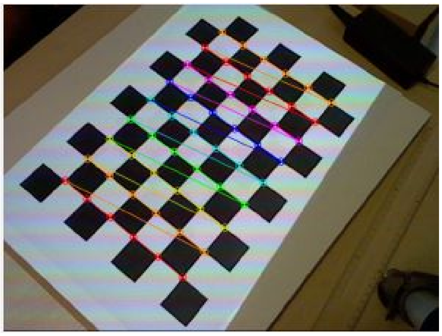


Figure 18: A chessboard typically used in camera calibration. The physical size of the squares are known and can be automatically detected using computer vision.

CALIBRATION The calibration process can be conducted in a number of ways. The majority of them are based or inspired by Zhang's method [194]. It estimates both the intrinsic and extrinsic parameters using a set of correspondences between 2D points located on the imaging sensor of the projector or camera with their positions in world space. Obviously, we could consider selecting points in the imaging plane and manually measure their location in world space. However, manual measurements are very tedious and error prone. Alternatively, we can consider a semi-automatic method to help us generate these correspondences more quickly.

In the case of a camera, you can use a predefined pattern which contains points of physically known positions that is also easily detectable using computer vision. A black and white chessboard is usually a good choice. Therefore, since we can automatically detect the chessboard corners in the image – as seen in Figure 18 –, correspondences between image plane positions and real-world positions can be generated quickly by simply moving the chessboard to different locations.

Calibrating a projector is similar. However, since the projector cannot "see", the problem has to be tackled in a slightly different way. First, a standard camera is calibrated so that it is possible to convert pixel positions in real-world distances. Then, instead of using a printed chessboard, we project one on a planar surface using the projector. Since we are projecting the chessboard ourselves, we already *know* the corners' positions on the image plane. Since the chessboard corners can be detected automatically by the camera and that their pose can be determined using the camera calibration, correspondences between the image plane of the projector and the real world coordinate system can be generated. Different tools and methods exist to calibrate a camera-projector pair together – for example, see [4, 108].

2.2.2 *Anamorphic Illusions*

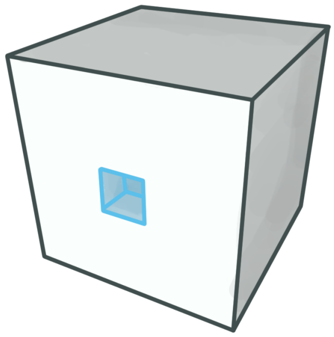
The process to generate anamorphic illusions is slightly different than texture mapping. However, it requires the same basic components: knowledge of the physical world's geometry, in real time, and the location of pixels in space – i.e. calibration. In addition, we also need to know the user's viewpoint. Texture mapping consists in projecting *flat* textures on the surface of real world objects. Therefore, the object will appear correctly augmented from any viewpoint. If we are interested in the creation of *3D illusions*, this is where we have to rely on anamorphosis. The process is very similar to the creation of a fishtank VR [175] experience, except that SAR gives more flexibility – the whole environment can act as the display.

To illustrate the process of creating such 3D illusions, we will take the example of a simple physical cube. We will create the illusion that, on one of its face, there is a cubic hole – as shown in Figure 19a. The cube is physical and the hole will be virtual (drawn in blue). The illusion can be generated using the following steps (which are all illustrated in Figure 19 and are indicated in parentheses):

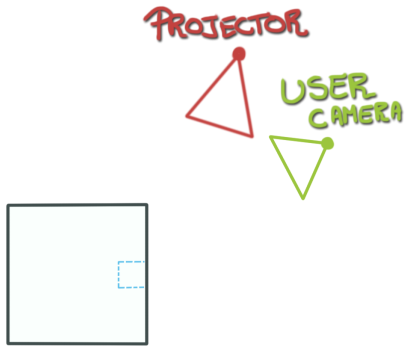
1. Create a virtual camera corresponding to the current user's point of view of the scene (19b).
2. Render the virtual elements (the small cubic hole only) of the scene from this virtual camera (19c).
3. Reproject the rendered image back onto the virtual replica of the real scene – a simple cube without any hole in it (19d).
4. Create a second virtual camera corresponding to the projector's point of view using extrinsic and intrinsic parameters retrieved from the projector calibration (19b).
5. Render the scene from this virtual camera (19e).
6. Project the rendered image on the real world environment with the projector which, from the point of view of the user, will appear correctly (19f).

2.3 TOOLS OF THE TRADE

This section gives a quick overview of the tools used for the different projects highlighted in the following chapters. Note that they could easily be swapped for other tools achieving the same purpose, but we will justify their choice briefly.

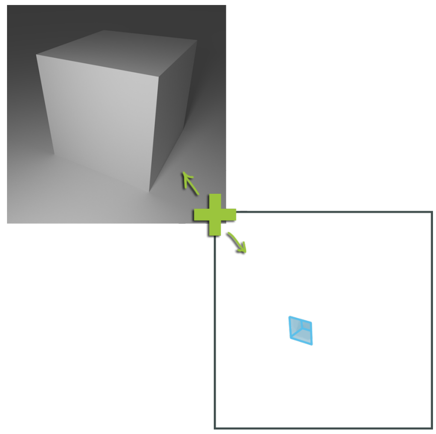


(a) Desired illusion



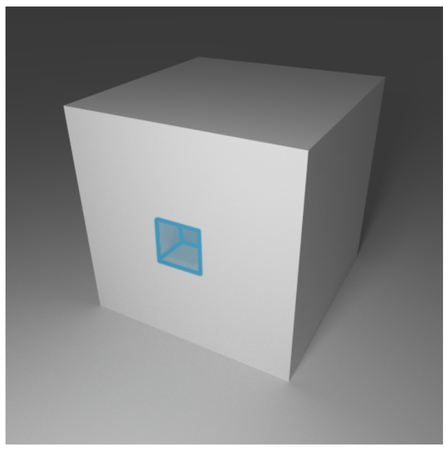
(b) Top view of virtual scene

(c) Rendered image from user's camera view



(d) Reprojection on virtual scene

(e) View of the projector.



(f) Final view for the user

Figure 19

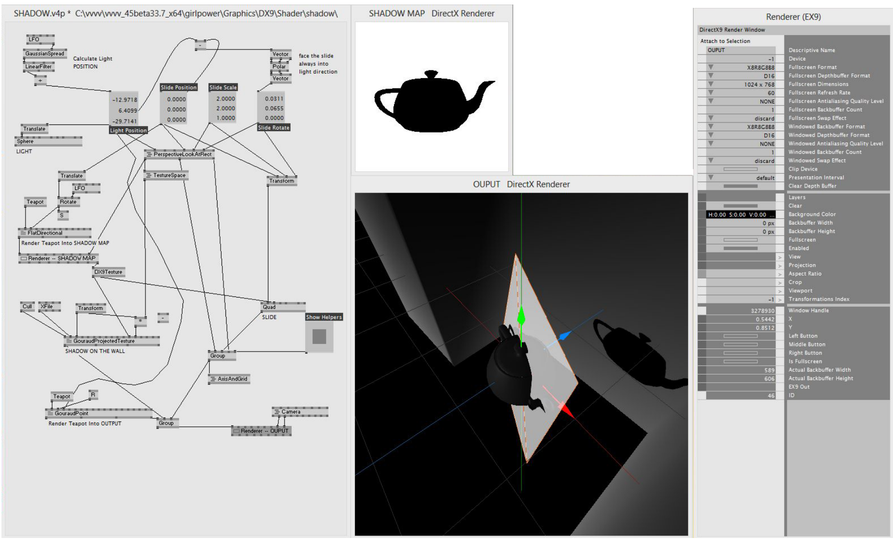


Figure 20: A view of the vvvv visual programming environment.

2.3.1 *Programming: vvvv*

vvvv[10] is a general purpose, visual programming toolkit. It is the backbone framework of all the projects described further. vvvv, while being “general purpose”, has a strong focus on visual installations – see Figure 20. There is only one mode: real-time. Everything is done *interactively*, which avoids the dichotomy of “code *then* debug” of standard integrated development environments (IDEs). Therefore, it is easy to experiment or make fixes in real-time which is especially helpful with systems dealing both with digital and real world components. Moreover, for interactive systems, being able to tweak parameters *while* interacting with the system proves invaluable. The framework is built on Microsoft DirectX 9/11 and plugins can be created, again in real-time, using the C# language. Shaders can also be written in HLSL, in the same way leveraging the immediate feedback properties of visual programming. Moreover, it natively supports using networks of computers for installations that require more processing power. From a “master” computer, it is possible to seamlessly assign certain parts of the program to slave machines. In a projection setup, for example, each projector can be assigned its own computer and have the processing be done “locally”. It is free to use for non-commercial and educational purposes.

vvvv is part of a movement called “creative coding”. It emphasize code as a mean of expression and of artistic creation. Usually, these frameworks try to present high-level APIs enabling their users to experiment and tinker quickly – a process essential to creative work as is highlighted in Victor’s talk [167]. Often, this take the form of very easy access to computer vision algorithms, interactive devices

10 http://vvvv.org/

such as the Microsoft Kinect and the Leap Motion and the easiness of controlling graphics pipelines which, in their basic forms, are really difficult to handle to novice programmers and artists. Examples of popular creative coding toolkits include Processing[11], openFrameworks[12], Pure Data [13], Max MSP[14] and TouchDesigner[15].

2.3.2 *Tracking: OptiTrack*

The tracking of the augmented objects was handled with OptiTrack cameras. As mentioned before, this tracking solution provides real-time performance and precision which is required when working with augmented objects of small size – often smaller than 10 centimeters.

2.3.3 *Calibration*

Calibrating an installation with an OptiTrack camera is slightly different than if the tracking would be provided by a standard RGB camera such as is the case for fiducial marker based tracking (Section 2.2.1.2). Indeed, OptiTrack cameras are IR cameras, which means that they do not "see" visible light. Therefore, it is impossible to capture the light projected from a standard projector and consequently, a projected chessboard required for calibration. Some camera models, however, have an IR filter switch, allowing to turn off the IR filter. However, the resolution of these cameras are relatively low – 640x480 pixels in the case of the V120:Trio model – making it difficult to obtain high-quality calibrations.

For these reasons, we calibrated the tracking system with the projector with a semi-manual method. Reflective markers were installed on a standard printed chessboard pattern (Figure 21). That way, using the pose of the rigid body returned by the tracking system, we are able to infer the world position of each chessboard corner. Then, we manually create corresponding point pairs by selecting points *using the projector*. That is, for each given chessboard corner expressed in world position, we select the corresponding point in the real world by looking at a projected cursor *onto the chessboard*.

2.4 WHAT IS SAR GOOD FOR?

Spatial augmented reality in the context of this thesis is a technological choice. It presents advantages as well as challenges both in terms

11 https://processing.org/
12 http://www.openframeworks.cc/
13 https://puredata.info/
14 https://cycling74.com/products/max/
15 https://www.derivative.ca/

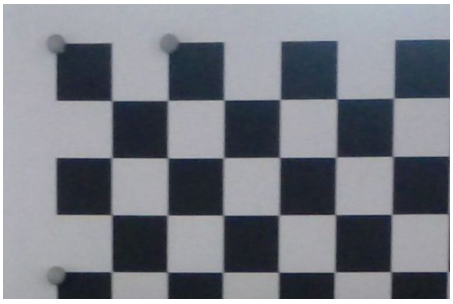


Figure 21: A chessboard pattern with reflective markers attached to it.

of technology and interaction. This section will first discuss the concept of "presence" which is important to any technology aiming at creating a seamless experience. Then, the benefits and drawbacks of SAR will be discussed.

2.4.1 *Presence*

Presence is defined as the subjective experience of being in an environment, even though one is not physically located in said environment [185]. It is associated to virtual reality which the whole purpose is often to immerse one or many persons in computer-generated artificial worlds and giving the illusion of "being there". However, one of the reasons we are interested in SAR is its potential to anchor the digital information *in the real world*. Therefore, presence, the way it is defined, is not appropriate for our context. For this reason, Stevens and his colleagues proposed the *object-presence* concept [146]. Inspired by the definition of presence given by Witmer and Singer [185], they define object-presence as "the subjective experience that a particular object exists in a user's environment, even when that object does not". Bennet and Stevens [17] then evaluated the effect of touching physical objects augmented with front projection. They found that directly touching – either with direct touch or using a TUI – an augmented object lowers object presence significantly. They hypothesize that it is due to the occlusion of the projection created by the hand – onto which projected content appears – and the fact that the weight and texture of the physical props does not correspond to the visual augmented representation. We will discuss this issue more in the following section 2.4.3.

2.4.2 *Benefits*

First and foremost, SAR presents the benefit of being anchored in the real world, which is a key component for our line of research. By displaying digital elements *onto* the reality directly, it has the potential to leverage all the others cues we are used to have in the physical realm. Moreover, it is worth mentioning its capabilities to enhance or complement physical artifacts which are physically *invaluable* and *unique*, such as the ones presented in museums – e.g. [10, 132]. Using projected light also has the benefit of being flexible in terms of the installation size. For example, it is possible to create dynamic rendering over huge buildings or very small objects with relatively similar setups[16].

When first introducing SAR, Raskar et al. [128] made a list of the different advantages of SAR in comparison to standard AR. Namely, they mention the fact that the user does not need to wear any device such as a HMD. A larger field of view can also be supported; using multiple projectors allows for covering an entire room with projected light. Also, since the virtual objects are displayed near their real world location, eye accommodation is made easier.

2.4.3 *Drawbacks*

While SAR comes with many advantages and flexibility, it suffers from different shortcomings. Already identified by Raskar et al. [128] upon the presentation of the SAR paradigm is the important reliance on the display surface properties. That is, projecting on highly specular or dark colored diffuse materials render the projected content almost invisible. Another mentioned problem is the shadow cast by a user manipulating an object. This problem breaks the illusion in two ways: it casts a shadow on the object and it displays the projected content onto the user's hand. The first problem can be tackled to an extent by increasing the number of projectors used to augment the scene. Indeed, with careful placement and blending, it is possible to avoid shadows to appear on the surface of the augmented surface [5]. Concerning the second issue, it should be relatively easy to create an occlusion mask to avoid projecting on the user occluding an augmented object. This could, for example, be achieved using the Kinect camera point cloud combined with its skeleton tracking capabilities. The overall issue with shadows can alternatively be worked around by using rear-projection system such as in [14]. Note, however, that rear-projection systems require a lot more space and constraints. Small objects would require having a pico projector embedded in them in a very stable manner. As of today, this is difficult to achieve.

Using many projectors also make installations more complex to install and calibrate in comparison to see-through or video see-through AR. Projectors are more complex to calibrate than cameras for the simple reason that they are output devices that are not equipped with sensing capabilities – as is explained in Section 2.2.1.3. More-

16 Of course, projecting over buildings often require more projectors and higher lumens output, but the techniques remain the same.

over, multi-projection installations require overlapping sections to be carefully blended together to create a uniform augmentation. Finally, SAR makes it difficult to display things in "mid-air" compared to other means of achieving AR. One of the way to achieve this is to resort to 3D illusions involving anamorphosis (as discussed in Section 2.2.2). Alternatively, it is possible to create a temporary mid-air diffusion medium on which to project, such as smoke [141] or mist[99]. Usually, such 3D illusions rely on stereoscopic projection such as the MirageTable installation of Benko et al. [15]. However, this has the main drawback of requiring that the user wears glasses[17], which is something we wanted to avoid in the first place.

2.5 SUMMARY

In this chapter, we provided a background for augmented reality and, more specifically, spatial augmented reality. The spatial aspect consists in displaying the augmented information in the environment itself, either relying on screens or projectors. We covered the main techniques used in SAR, which are texture mapping and anamorphic illusions. We also presented an overview of the working principles and ways to achieve them. The notions presented here served as a basis in creating most of the systems that will be covered in Part ii and iii of this thesis. We will refer back to appropriate sections of this chapter later on.

17 In addition to the other drawbacks related to the use of stereoscopy such as discomfort, fatigue, stereoanomaly, etc [121, 84].

Part II

INTERACTION

When combining real objects with digital content displayed onto them, interacting with the digital content is not trivial. How to interact with a medium that is virtual but that is hosted on reachable and manipulable physical objects? We investigate the possibility to use traditional interaction devices and computer environment – the desktop – as a starting point. We first investigate if it is possible to interact with augmented objects with such devices, even without the presence of a screen. Then, we propose a system leveraging both the use of screens and traditional input devices in combination with tangible augmented objects to allow the creation and customization of content for these objects.

3

CURSAR: POINTING IN SAR FROM 2D POINTING DEVICES

Figure 22: A sketch describing a standard desktop environment which would allow a mouse pointer to travel out of the screen and into the surrounding environment to acquire targets.

In the previous chapters, we have discussed the idea that augmented physical objects could be an interesting way to mesh together digital capabilities while remaining anchored in the physical world, which our bodies and minds inhabit every days. However, this raises the question on how content for these mixed reality objects will be created and *interacted* with. While technical issues related to the augmentation of objects are being solved progressively, the problems related to interaction remain largely unexplored. A central concept in interaction is the ability to select an element or position. In this chapter, we investigate the question of pointing – more specifically using standard 2D pointing devices, as shown in Figure 22 – in a SAR context.

3.1 CONTEXT

There are many components to interaction. A fundamental one consists in pointing. Several strategies to point at augmented objects exist. When the augmented content is visualized using a multitouch device, via video see-through, it is possible to leverage the touch surface to point at a target, as is done in Vincent et al.'s work [168]. Regarding pointing on *tangible* augmented objects, one option is to touch directly the area of interest. Technological means to achieve precise touch detection on physical objects are becoming mature and versatile. For example, the field of OUIs – see Section 1.3.4 for an overview – envisions everyday objects covered by thin film displays which are multitouch. PapARt is a system which detects touch on standard tracked sheets of paper [86]. Touché is a system that enables the detection of different types of touch – one finger, two fingers, palm, knuckles, etc – on almost any object [137]. OmniTouch uses a depth sensor to add touch capabilities on any surface [50].

Direct touch, as an approach, is very straightforward and, consequently, it may be valuable in many contexts. However, it suffers from many drawbacks. In particular, anatomical issues such as the "fat finger" problem and the fatigue that is linked to mid-air interaction make direct touch little adapted as soon as accurate and prolonged actions are required (e.g. professional object design). In addition, direct touch is not possible when dealing with very fragile objects (e.g. relics in museums) or as soon as the objects are out of reach. For distant interaction, laser pointers or virtual rays can be good alternatives, but they still suffer from similar accuracy and fatigue issues. This motivated our approach to consider the use of standard pointing devices, namely mice and graphics tablets, to point at augmented objects. One of the thing we had in mind is the creation of hybrid applications – as shown in Figure 22. These applications could leverage both the power and flexibility of professional tools, currently hosted on traditional computing platforms, and the real world surrounding these platforms. Years of human-computer interaction (HCI) have shown that mice and graphic tablets are decidedly well suited to point at visual objects displayed on 2D screens. Our assumption is that they can benefit to SAR as well, as soon as precision or prolonged work is required.

As an example, we can imagine an inspection scenario where an engineer points at an augmented circuit board with a mouse to highlight defects on small components. Another example is a design scenario where the artist draws by way of a graphics tablet on a physical object, e.g. a 3D-printed one, to give it a specific appearance. For these two scenarios, it is interesting to note that the user equipped with a standard pointing device is still able to interact efficiently with standard GUI components displayed on a traditional screen, opening the way to true hybrid applications.

Pointing from mice and tablets has been extensively studied in traditional HCI contexts. In particular, Fitts' law [96] is able to predict the speed at which a user will be able to select a target depending on its distance and its size. Other works have been dedicated to pointing in 3D stereoscopic contexts [138, 154]. The current work is the first one that studies the question of pointing in SAR, from standard pointing devices. In this work, we are interested in a setup where the user is sitting at a desk (desktop environment) and is interacting

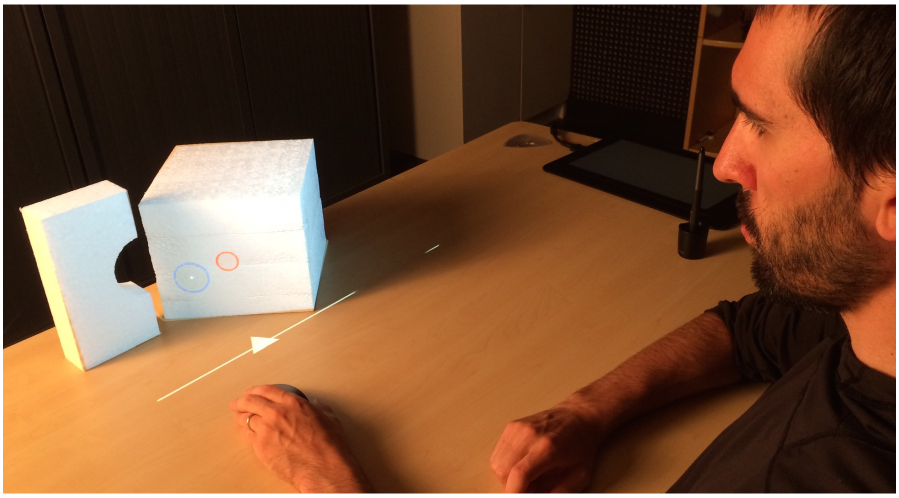


Figure 23: A user moving a cursor (represented in blue) to a target (represented in red) on an augmented object by way of a standard mouse.

with objects located in front of him or her, as shown in Figure 23. Our contribution is the evaluation of the performance of pointing in a SAR environment using a standard pointing device compared to a traditional screen-based setup.

3.2 RELATED WORK

Some research projects explored interaction with projected content, in a SAR context. Bandyopadhyay et al. [9] proposed the first interactive SAR prototype allowing users to "paint" physical objects with projected light using a six degrees of freedom tracked stylus. Physical-virtual tools [98] is a refinement of this concept, introducing more flexible edition tools inspired by real physical tools (e.g. an airbrush). Benko et al. [15] interacted with stereoscopic SAR using a mix of tangibles and gestures. These systems aimed for interaction modalities close to real-world metaphors. However, while perhaps more natural, they might prove to be less suited for precise and prolonged work than traditional 2D input devices.

The concept of pointing in SAR is similar to pointing in other contexts, namely multi-display environments (MDEs) and stereoscopic displays. In some ways, SAR can be compared to MDEs in that the physical world acts like a continuous space comprised of small display surfaces. As with MDEs, SAR might have some blind spots where the cursor will disappear because of a lack of projection support. Mouse Ether [11] and Perspective Cursor [110] are both systems that were developed to circumvent problems related to switching from one screen to another. The work of Xiao et al. [190] consists in projecting a cursor that can slide on any surface of the environment (which has been modeled in 3D beforehand). However, the system has been designed as a way to give feedback on the cursor's position when *transitioning* between screens and no targets were located in the environment itself in their evaluation. Pointing on a stereoscopic display has been studied by Teather and Stuerzlinger [153, 154]. They studied different cursor types in what is effectively a "2.5D", or projected pointing task, using a 3D Fitts' law pointing task. We also used projected pointing in this study. However, working on a real-world canvas is different from working on a screen since the real-world does not provide any reference frame for the 2D interaction. Moreover, a SAR installation does not suffer from the vergence-accommodation conflict present when using stereoscopic screens. Closer to a SAR setup, Reikimoto and Saitoh [130] proposed a spatially continuous workspace, allowing users to drag and drop content across different surfaces and objects. However, the pointing activity was not studied.

3.3 POINTING IN SAR

Our SAR environment is comprised of a static scene laid out on a table in front of the user. A projector is then used to augment the objects.

On a standard screen configuration, the mouse cursor is generally represented as an arrow moving on the screen plane. When a 3D scene is displayed, the user is able to select any visible part of this virtual scene by picking the rendered result at the cursor location (Figure 24). Since most people are already experienced with this way of pointing, we wanted to know if this technique could be ported to a SAR environment albeit the lack of a physical screen support. Therefore, we used exactly the same metaphor in SAR, with the difference that the 3D scene is physically there, while the screen plane becomes virtual. The user moves the cursor on this virtual plane as he or she would do with a physical screen, as illustrated in Figure 25. A line representing the intersection between the virtual plane and the table is projected onto the table, and an arrow indicates the horizontal position of the cursor – see Figure 23 and 25 (item C). Contrary to standard screen configurations where the cursor is displayed on the screen plane, our SAR cursor is displayed directly on the physical objects. This cursor is represented as a cross within a 2D circle that is aligned with the underlying surface. Technically, we cast a ray formed by the eye and cursor position on the virtual plane towards the scene. We then position the cursor perpendicularly to the normal of the picked point. The visual feedback (line and arrows) helps to know where the cursor is as soon as the latter does not project onto an object.

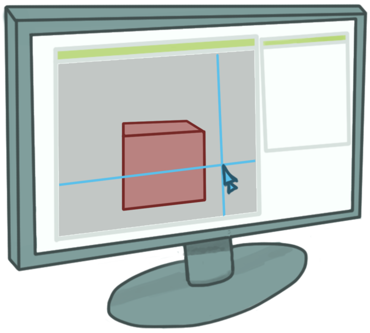


Figure 24: The standard way to point on a 3D scene displayed on a screen. The cursor moves in the window plane and a virtual point is selected by picking the rendered scene at the cursor location.

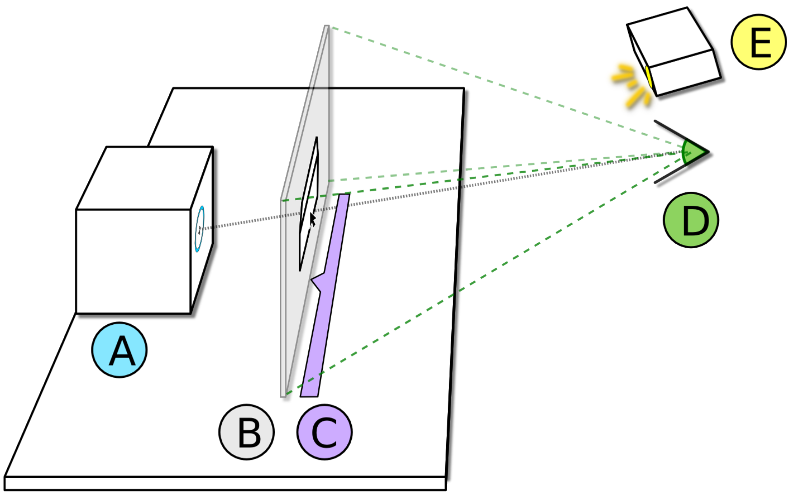


Figure 25: Drawing of the experimental setup. (A) Objects composing the scene to be augmented on which the cursor is displayed (light blue halo). (B) Plane on which the cursor is projected. This plane is either virtual in the SAR condition or physical (white wooden panel) in the SCREEN condition. (C) Feedback used in the SAR condition indicating the position of the virtual plane with the tip of the triangle indicating the horizontal position of the cursor. (D) The position at which the user is viewing the scene. (E) Projector.

3.4 USER STUDY

We conducted a user study to assess the performance of the pointing technique described in the previous section (SAR) in comparison to a screen-based baseline (SCREEN). Our research question was the following: What is the difference in performance of a pointing task realized on a screen compared to one realized with a SAR installation given that all other conditions are constant? Does pointing in SAR follows Fitts' law?

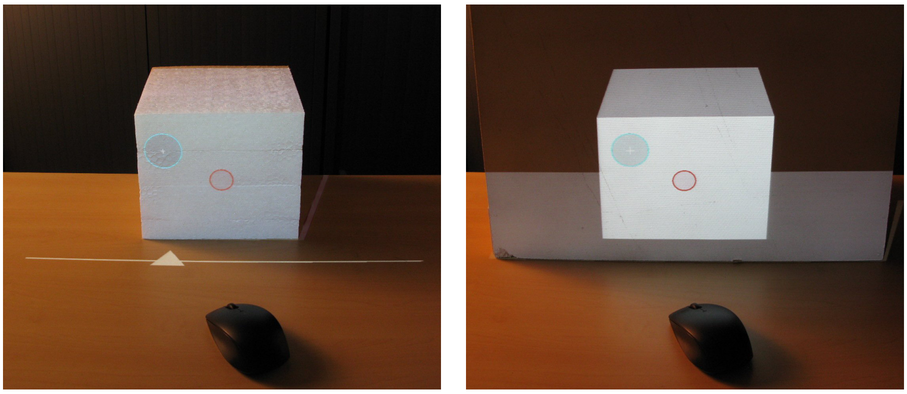
3.4.1 *Participants*

Sixteen participants took part in the study (12 males, 4 females, mean age 28.75, SD 4.71). All of them obtained a university degree. Six participants were left-handed (the mouse used during the experiment was adapted to both left- and right-handed users). All the participants were familiar with mice, whereas they had very little experience with graphic tablets. None of them had previous experience with SAR systems.

3.4.2 *Apparatus*

The scene to be augmented was laid out on a table in front of the user. Each object of the scene was manually measured and modeled in 3D. A projector was located above and behind the user pointing at the scene. The projector was calibrated using OpenCV's camera calibration functions. We used a 3.6 GHz Core i7 PC with Windows 8 equipped with two GeForce GTX690 graphic boards. The video projector was a ViewSonic Pro9000 with a resolution of 1920×1080 pixels. The same setup was used for both SAR and SCREEN conditions. In SAR, the virtual scene was projected directly onto the physical objects whereas a white wooden surface located at the same position was used in the SCREEN condition. This ensured a similar frame rate (50 FPS), colorimetric configuration (color, brightness, contrast) and approximately same pixel size in both conditions. The focus of the video projector was set on the screen plane. On this plane, the resolution was effectively of 915×904 pixels.

In the SAR condition, the objects were augmented by reprojecting the virtual scene from the point of view of the projector. In the SCREEN condition, the viewpoint of the user on the scene was virtually reproduced and reprojected on the virtual counterpart of the physical screen. Then, this reprojection was rendered from the point-of-view of the projector, effectively making the viewed scene in both conditions identical, as shown in Figure 26. We did not use real-time head tracking, but the user head's position was measured manually



(a) SAR condition (b) SCREEN condition

Figure 26: Comparison between the SAR and SCREEN conditions, captured from the same point of view. The SCREEN condition is made to be a replica of the view of the reality – note the virtual and real table which are aligned.

and thus accounted for. The whole installation has been created using the creative coding framework vvvv.

For the input devices, we used both a mouse (MOUSE) and a Wacom Cintiq 13HD tablet (TABLET). The screen of the tablet was not used for the experiment and, therefore, was displaying a black viewport. The button located on the pen was used for the selection action. The mouse was used in a relative mode while an absolute mapping was associated with the tablet. The acceleration transfer function of the mouse was disabled.

The 3D scene was composed of a 21×18×21 cm cube, as well as a more complex shape with comparable dimensions (see Figure 23). The scene onto which the targets to acquire were laid out varied by rotating the cube by an angle of 45° to provide more depth changes between trials. The participants sat at a distance of 1 m from the screen or physical objects, and the height of the chair was set in order for the participants' head to be located at the ideal observer position.

3.4.3 *Procedure*

We followed the procedure described in MacKenzie's work [96]. The participants had first to position the cursor in a home area represented by a red circle. After one second, this circle moved from red to green and a target appeared in the scene. The participants were instructed to select this target as quickly and accurately as possible. The start time was recorded when the cursor left the home area and stopped when the users clicked on the target. The targets were spread on a circle centered on the home area.

FACTORS	TIME (ms)	INEFFICIENCY	ERRORS	THROUGHPUT (bits/sec)
Input: Mouse/Tablet	ns	0.16/0.22 (SD: 0.05/0.08)*	ns	ns
Output: Screen/SAR	846/959 (SD: 154/119)**	0.17/0.21 (SD: 0.07/0.07)·	ns	5.75/3.84**
Grand average	902 (SD: 404)	0.19 (SD: 0.30)	0.05 (SD:0)	—

Table 1: Statistical results. Marks: ** for $p < 0.01$, * for $p < 0.05$, · for $p < 0.1$, ns: not significant, —: not applicable.

3.4.4 *Design*

We used a 2×2 within-subjects design. The independent variables were the output modality (SCREEN, SAR) and the input modality (MOUSE, TABLET). The dependent variables were the completion time, the inefficiency defined as $\frac{Path_{actual} - Path_{optimal}}{Path_{optimal}}$ [193] and the number of errors, defined as the number of selections outside the target area. For each condition, the participants had to acquire 40 targets, resulting in 160 target acquisitions per participant, and 2560 records in total. The order for the input and output were counterbalanced following a latin square to avoid any learning effects.

3.5 RESULTS AND DISCUSSION

Because the homogeneity of variance couldn't be verified according to Levene's test ($p < 0.001$), we analyzed our data with non-parametric statistics, using multiple Wilcoxon signed-rank tests and false rate discovery correction. We retained trials which did not comprise errors to study time and inefficiency across our factors. Statistical results are reported in Table 1.

TIME There was no significant effect of the input device on completion time. However, output modality had a significant impact. Users were 11% faster in the SCREEN condition compared to the SAR condition. While having higher completion time, the drop in performance is relatively low, especially considering that the cursor reference frame was virtual.

INEFFICIENCY Inefficiency is a measure of "wasted" cursor movement by the user. Input modality had a significant effect, the tablet being more inefficient than the mouse. This difference can be explained by the lack of experience of almost all participants with such a tablet. Output did not have a clear significant effect on the inefficiency of the movements of the users. When looking at Figure 27 one

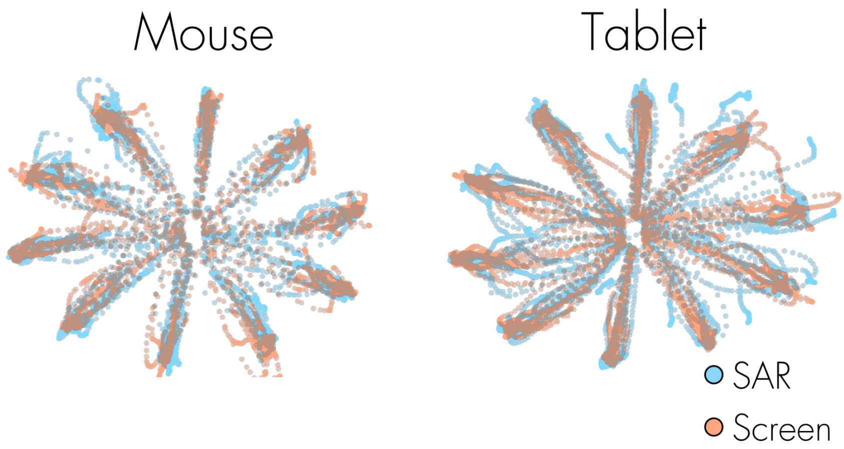


Figure 27: Trajectories of all the trials in the scene with maximum depth changes. The center is the starting point of the trials and the tips are the location of the targets.

can observe that the participants tend to follow similar strategies for reaching the target. In both cases, straight lines are drawn, independently of the underlying background (2D flat screen or physical 3D objects).

ERROR RATE There was no significant effect of either input modality or output on the error rate. On average, the error rate was 5%.

THROUGHPUT The target condition is reflected by the Index of Difficulty (ID), which indicates the overall pointing task difficulty. $ID = log_2(\frac{D}{W} + 1)$ [96]. D is the projected target distance in the virtual screen and W is the perceived target size. W varied according to the location and orientation of the target in the scene. ID was discretised from $[1.91; 4.92]$ to $[2; 5]$ by steps of 0.5. We averaged the completion time across ID and conditions (input$\times$output). We modeled the movement time (MT) with a linear regression. We obtained an adjusted R^2 value of 0.8479 which shows that the completion time of pointing tasks in SAR using mice and tablets still follows the Fitts' law (see Figure 28), and consequently remains predictable. We also computed associated measures of performance, also known as "throughput", using the slope of the regression lines. $Throughput = \frac{1}{b}$ [193].

There was no significant effect of the input device on the throughput, whereas output device did have an effect. The screen condition was significantly more efficient than the SAR condition although, as it was the case for the completion time, the difference is relatively low.

Overall, the participants were slightly less efficient in the SAR condition than the SCREEN one. This difference could be explained

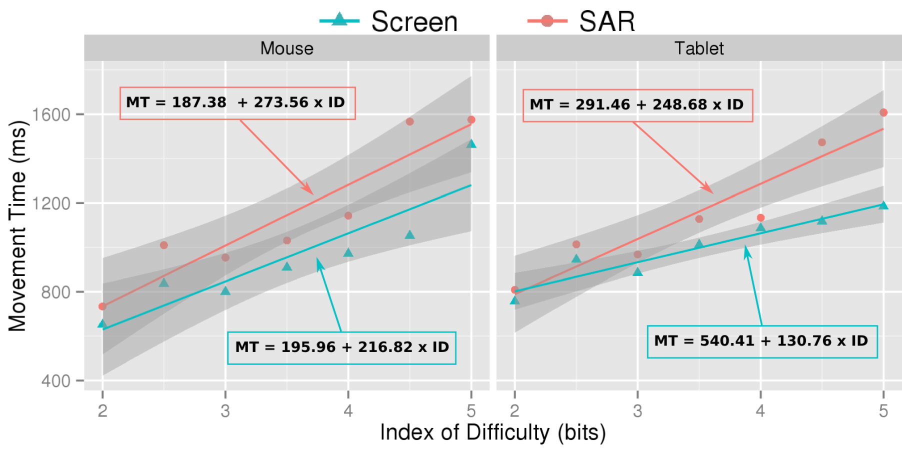


Figure 28: Fitts' law models. $R^2 = 0.8479$.

by the years of experience of the participants with pointing in front of a screen whereas they were exposed to a SAR setup for the first time. Also, it is interesting that removing the physical reference frame (screen) of the cursor does not prevent users to interact in the same way they are used to, i.e. as if a physical screen was there. We can thus presume that with additional experience, participants may improve their performance with SAR. Another possible cause for the drop of performance is the presence of blind spots where the cursor disappear because of a lack of projection support (such zones were involved in about 1/4 of the trials). It could be interesting to compare the effect of these gaps in MDEs vs SAR to evaluate the impact of the frame of reference provided by the screen. Additionally, possible extensions of this work include studying the performance when moving the viewpoint of the user while using the Perspective Cursor [110] and evaluating if the performance drop observed in the SAR condition can be reproduced with other interaction techniques such as laser-pointer.

3.6 SUMMARY

We presented an approach for interacting with desktop SAR, i.e. when the user interacts with physical objects in front of him/her by way of standard pointing devices. A user study has shown that Fitts' law remains valid even if no physical screen is present. Users are able to point at targets displayed on the augmented objects in a manner that is comparable to what they used to do in front of a standard screen. This finding opens interesting perspectives, allowing desktop SAR applications to be used to extend the current desktop setup with augmented physical objects. This is the topic we will investigate in the next chapter.

4 BRIDGING DESKTOP COMPUTERS AND TANGIBLE AUGMENTED OBJECTS

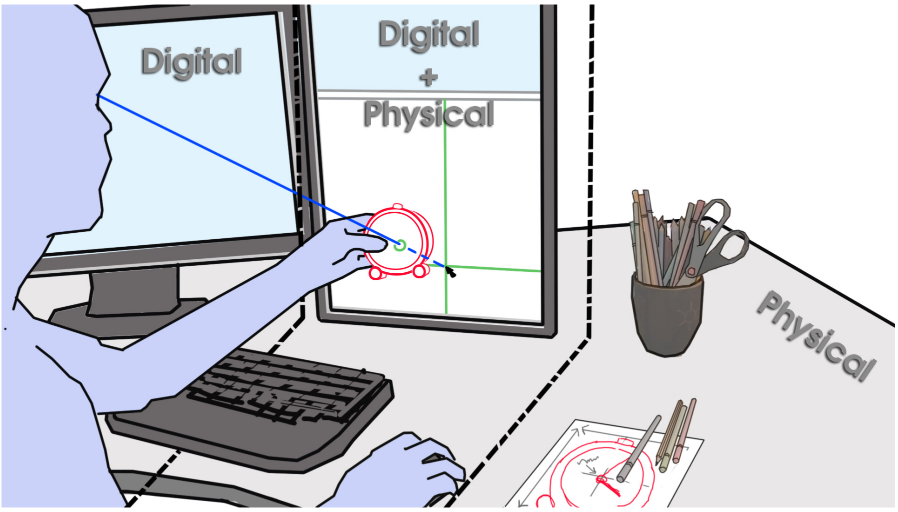


Figure 29: A workspace that allows a user to work both digitally (left) and physically (right). A tangible viewport (middle) allows physical objects to be used in the context of a screen.

This chapter presents Tangible Viewports (Figure 29), a proof-of-concept prototype that builds on the ideas presented in Chapter 3 to create digital content on physical objects that are augmented using SAR. With CurSAR, we studied the performance of a standard pointing technique using 2D indirect devices, like the mouse, to reach targets displayed in the real world using SAR. We found out that pointing in SAR remains fully usable with a minor performance drop (11%) compared to a screen condition. However, performance still is impacted negatively when removing the screen context. One of the research objectives tackled in this thesis is the digital content creation and interaction for tangible augmented objects. Since most digital content creation tools are designed for standard input devices, we combined traditional computer screens, physical objects and the pointing technique studied with CurSAR in a single system. Effectively, it enables a seamless integration between desktop applications and physical objects. The objects behave in the same way as they would if they would be rendered in a traditional viewport on screen.

4.1 CONTEXT

Look at your workspace right now. There is a high probability that it is divided into two different areas: one for working digitally (computer) and one for working physically (pen and paper, books, building materials). Figures 29 and 30 highlight such a typical desk. This dichotomy has been present in our work environments for a long time, and a lot of effort of the Tangible User Interfaces (TUI) community has been directed towards a digitally enriched physical space. Compared to the traditional mouse-based paradigm of computers, tangible interaction [65] has been shown to provide richer interaction experiences that are especially well suited for collaboration, situatedness and tangible thinking [142].

4.1.1 *Tool-Based Interaction with SAR*

Many different works are inspired by the use of physical tools for interaction with digital content and for its creation. Touch Tools [51] leverage the muscle memory of our hands when using physical tools such as holding a pen or a camera to trigger the proper behavior on a multitouch surface. IntuPaint [162] and MAI painting brush [118] are both systems that use a physical paint brush to digitally paint real objects while at the same time keeping the feeling of using a real brush. In the same line of thought, Conté [169] is a tangible tool inspired by the multiple ways an artist uses a drawing instrument such as the crayon.

Many SAR systems rely on TUIs for interacting with the digital content. They include tools for editing the appearance of physical objects [9, 98], sculpting [124, 97] and educational purposes [38]. TUIs are especially well suited for collaborative tasks and provide a strong situatedness [142]. The previously mentioned approaches are deeply rooted in physicality and stay away from traditional computer environments.

However, even when tangibility holds great promises for interaction, its use in real-world contexts remains rare, while we still use standard computers for the majority of our daily tasks involving digital information. The desktop computer is still a relevant tool to work with digital and physical matter. However, we also think that its place on our desks should be rethought [127, 130]. Instead of being considered as a self-contained platform that happens to be installed on a desk and its reach limited to the extent of its screen, it should be considered as a tool part of the whole toolset laid onto the desk, aware and capable of interacting with its surroundings.

4.1.2 *Augmented and Smart Spaces*

Augmented and smart spaces are systems that use see-through augmented reality or projectors, often in an office environment, to enhance the workspace. Raskar et al. [127] proposed the idea of a hybrid workspace that would combine the physical environment with a spatially augmented display system in order to create a continuous mixed-reality space. Similarly, Augmented Surfaces [130] is a system that creates interactive surfaces on a table, wall and laptop using projectors. Users could use their mouse cursor to drag information between the different surfaces. More related to desktop systems, Kane et al. [74] present a hybrid laptop-tabletop system that uses two pico-projectors mounted to a laptop computer to add interactive areas on the table around the device. The system is able to detect tangible objects on the table but does not augment them in any way. HoloDesk [56] is a situated see-through display where virtual and tangible objects can be manipulated directly with the users' hands, but does not integrate any traditional computer-related tasks. From an interaction point of view, the work of Lee et al. [89] is close to ours. They present a see-through desktop environment that supports transitioning from 2D and spatial 3D interactions easily. The system allows users to see the content of the screen and their hands behind it at the same time. Their main focus was on handling virtual elements. We are instead interested in bringing interaction with physical objects to a traditional desktop workspace.

There are different frameworks for interaction leveraging the use of SAR [163, 145, 191]. Especially worth mentioning, and closer to the system described in this chapter, Akoaka et al. [1] created a platform for designing interactive augmented objects using *either* natural interaction *or* a standard desktop computer. We instead merge both together.

4.1.3 *Combining Physical Objects and Desktop Computers*

We propose to leverage the potential of tangible interaction, while relying on the efficiency of standard desktop environments, in an integrated way. We present Tangible Viewports, a screen-based tool enabling the use of tangible objects in a standard desktop-based workflow. Contrary to many tangible user interfaces where tokens and generic props are used, we use the physical objects as canvases that are also the results of the creative or visualization process. These objects are augmented with Spatial Augmented Reality (SAR). SAR is especially well suited for creating a hybrid work environment where digital workspaces can be combined with physical tools and canvases [127, 130]. It is possible to interact with these objects through direct touch or using tools. When held in front of the computer screen, the

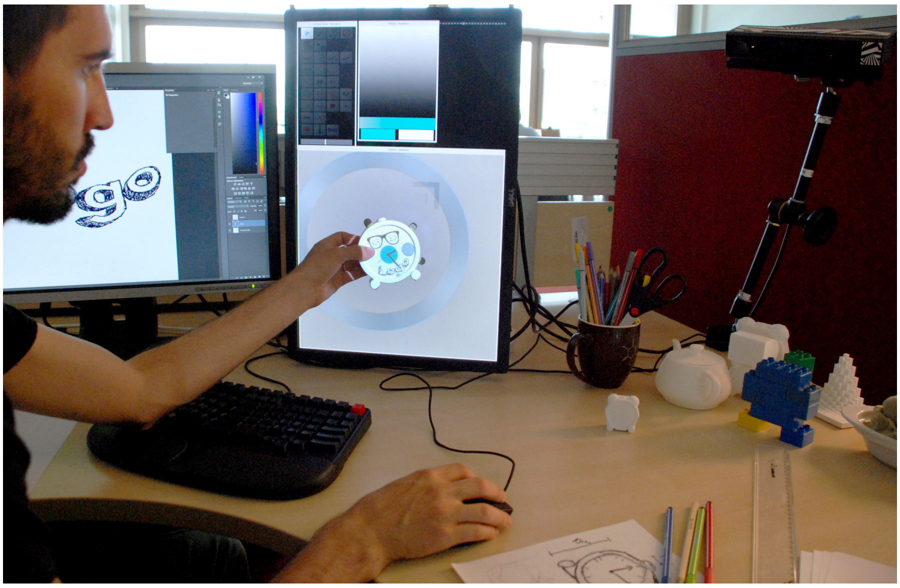


Figure 30: A real object can be placed in front of the screen to use it as a canvas for digital applications. Digital content is added using SAR.

mouse cursor can seamlessly slide from the screen *onto* the surface of the objects and interaction with native desktop applications becomes possible. For example, one can use a painting software to paint over the surface of the object as if it was part of the screen using the mouse cursor. From the viewpoint of the user, the object behaves just as a 3D model would when rendered in a viewport on the screen with the major exception that he can i) observe the object from a different viewpoint by moving the head and ii) reach out to grab the object with his hands and manipulate it freely as illustrated in Figure 30.

In this work, we emphasis the use of the desktop computer screen and its relation to augmented physical objects (Figure 30). This relation has been little explored as a complementary approach to tangible tools (e.g. [98]). We suggest that it can be leveraged to create true hybrid applications that reduce the gap between highly flexible and expressive software, currently trapped inside a flat rectangular screen, and the intuitiveness and graspable nature of our environment.

The main contributions of this work are 1) Tangible Viewports, an on-screen window that enables interaction between a desktop computer and a physical object located in front of it, 2) a proof-of-concept prototype of an integrated workspace that combines augmented physical objects and native applications, and 3) the exploration of the interaction space of this hybrid work environment.

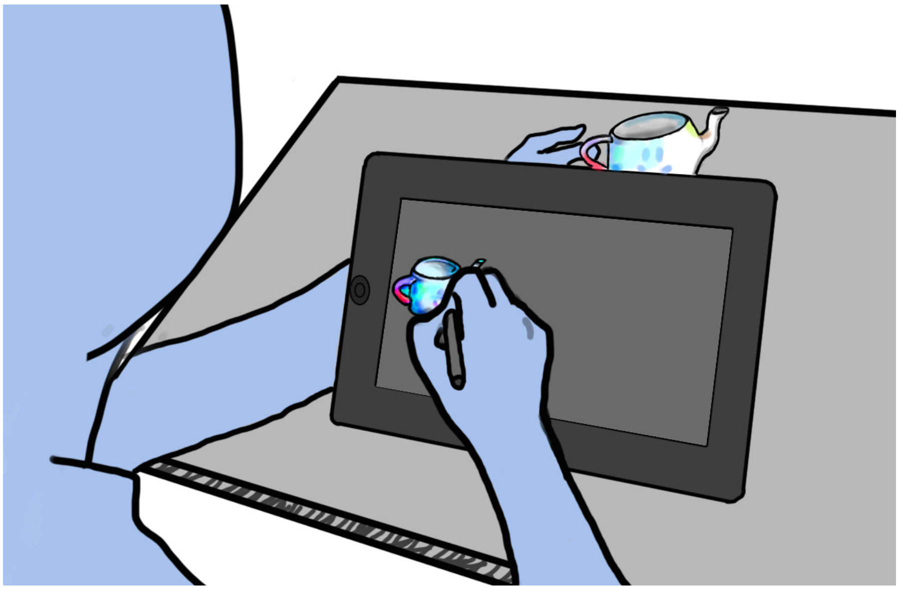


Figure 31: An augmented object is located behind a graphics tablet on which the user's view of the object is replicated. The user can draw on the object using the tablet. The drawings are synchronized on the physical object in real-time.

4.2 CREATING A SEAMLESS HYBRID SPACE

This section describes the design considerations and implementation details of the system. More specifically, we will discuss the screen position in relation to the physical objects and give details about the augmentation setup, the behavior of the cursor and how it is handled as well as the use of direct touch and tools to interact with the augmented objects.

4.2.1 *Screen Positioning*

The general idea of our system is to embed physical objects within the standard desktop paradigm. We consider the screen and chosen physical objects on the desk as spatial canvases where digital information can be displayed. This design differs from other approaches (e.g. [130]) that extend the reach of the cursor to the environment. We instead bring the physical objects within reach of the screen cursor. We tested different ways to position the screen in relation to the augmented objects before settling for the current design: in front, on the side and behind. We ended up choosing to position the screen *behind* the object.

4.2.1.1 *Screen in front object*

In this configuration, we positioned the augmented object *behind* the screen – see Figure 31. The screen was a graphics tablet replicating the user's view on the physical object. It was possible for the user to draw on the screen to digitally paint the object. The digital paint was reprojected in real-time on the physical object in the back. Also, both the user's head and the physical object were tracked which enabled the user to manipulate his viewpoint freely – the viewpoint is updated for both the graphics tablet and the real view.

The idea with this design was to ease the activity of drawing on an object by providing a flat surface onto which to draw – which is often easier to use for stability and flexibility (e.g. zooming on the drawing) purposes – while retaining the haptic properties of having a real object. Moreover, the synchronized views would allow for reviewing the design using the real object instead of a virtual rendering. The ability to handle the object with the hands also provided a good option to quickly place the object in the desired viewpoint. However, our first tests with this screen position encouraged the use of the physical object, not as a canvas, but simply as a proxy to control the orientation of the virtual model in 6 DoF. We instead wanted to put emphasis on the physical object itself because, even though the 3D virtual rendering often looks "nicer", the result of the object design process will evolve in the real physical world.

4.2.1.2 *Screen beside object*

An alternate configuration is to use the reality itself as an "extended desktop". This can be seen as a mix of multi-display environments and the work presented in Chapter 3 with CurSAR – an illustration of this configuration is shown in Figure 32. It consists in having a main display, a screen, where all the standard digital operations take place. When the cursor reaches the edge of the screen, it can seamlessly slide on the neighboring surfaces and geometries, such as in Reikimoto and Saito's work [130].

This design is interesting as it keeps a familiar interaction context with the use of indirect 2D pointing methods. That is, pointing with a mouse inside a screen is familiar since the screen provides a physical 2D plane on which the cursor is constrained. When moving outside the screen, the physical screen still acts like a reference system for the interaction scheme, which in some casual tests, seemed to prove useful. However, as demonstrated in the study of the previous chapter, pointing without a screen still has some costs in terms of performance. Moreover, it also has the drawback of requiring a context change when the transitioning from a 2D cursor located inside the screen to a cursor following the 3D geometry of the surrounding physical environment.

Figure 32: An augmented object is located *beside* a screen. When the cursor reaches the edge of the screen, it "leaks" out in the enviroment.

4.2.1.3 *Screen behind object*

Finally, we considered using the screen as a backdrop for the physical objects that lay in front of it, as illustrated in Figure 33. This configuration reduces the change of context required compared to having the screen beside the object – as discussed in Section 4.2.1.2. This design choice is supported by studies that have shown the very low performance drop for focal depth changes compared to angular movements [151, 24]. Moreover, in our current system configuration, normal use causes shallow depth of scene $-< 50cm$, a working space created by a typically recommended distance between the user and the screen – and users are not required to rotate the head position. When the screen is behind the physical object, the latter now appears to belong to the screen's interaction space. This is the configuration that was used to create Tangible Viewports.

4.2.2 *Spatial Augmented Reality Setup*

Our SAR setup is comprised of an augmented desktop environment and physical objects that can be brought in front of the screen. The objects can be manipulated freely by the user, or they can be placed on a support for convenience. Figure 34 illustrates the setup. The projector handling the augmentation is located behind the user and oriented so that its vertical field of view would span from the edge of the desk up to the top of the screen. It only emits light towards the physical object, so it does not perturb the visualization of the screen. The augmentation is generated using the techniques detailed in Section 2.2.

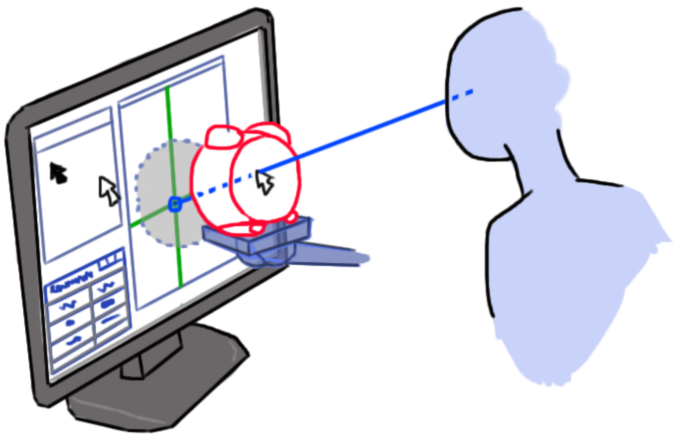


Figure 33: The user can interact with the augmented object when located in front of the screen, emulating the behavior of a normal viewport.

We use 3D printed objects created using a MakerBot Replicator 2 in white PLA plastic with a precision of ± 0.2mm. Alternatively, we could use already existing or sculpted objects given that they would require 3D scanning before, using KinectFusion [111] for example.

4.2.3 *Cursor Handling in Tangible Viewports*

The key element on which our system relies is the illusion that a physical object is entirely part of the screen space when located in front of it. In order to do so, we ensure that the cursor movements inside the working area occur in a continuous way, independently of where this cursor is displayed (screen or tangible viewport). The user thus perceives the visual space as a whole.

A window dedicated to the interaction with the object is created on the screen and its position is retrieved using the Microsoft Windows API. The screen is also tracked in world space by the OptiTrack system. Thus, knowing the 2D cursor position in the viewport space allows us to infer its position in world coordinates. A virtual camera is created to reproduce the user's view of the window (and whichever augmented object located in front of it). The user's head position is obtained using the Kinect v2 skeleton tracking. As soon as a physical object starts occluding the screen's cursor, from the observer's point of view, a 3D cursor appears at the correct location on the object, as illustrated in Figure 35. This is done by raycasting in world-space over the virtual scene from the user's viewpoint to the screen's cursor position (Figure 33). We thus obtain the 3D position and orientation on the first element on the line of sight of the user. The resulting transformation is then applied to the 3D cursor, which is displayed as a small disk aligned with the local surface's normal. This cursor

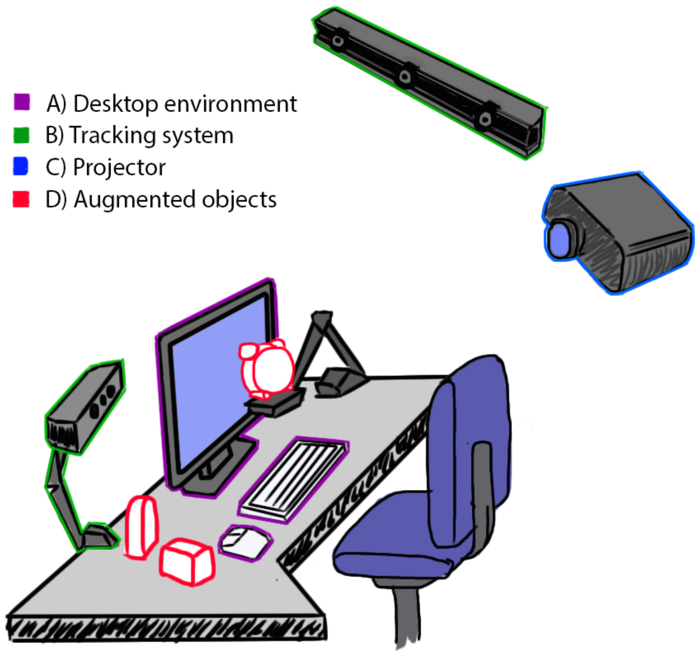


Figure 34: The SAR installation: A) The desktop environment using a standard screen and input devices, B) 6DOF tracking system (OptiTrack Trio and Microsoft Kinect v2), C) Projector and D) Physical objects that are being augmented.

is rendered as part of the virtual scene and reprojected onto the augmented object. On the screen, a horizontal and a vertical line passing through the cursor position are displayed in order to emphasize the link between the tangible viewport and the screen.

In the end, this technique is fully transparent to the users. They work with Tangible Viewports as they would do with any standard application. It is also to be noted that the head position of the user only impacts the *behavior* of the cursor; the cursor's appearance and the augmentations on the object are completely viewpoint independent. This is especially important for collaborative settings.

4.2.4 *Direct Touch and Gestures*

Beyond cursor interaction in front of the screen, direct touch on the objects is also supported. This is achieved by attaching a small reflective marker to a ring on the user's finger or on a tool (e.g. pen) so that it is detected by the OptiTrack system. We also tested the use of the Leap Motion in order to avoid instrumenting the finger of the user. However, the Leap does not support direct touch detection and is better suited for fine gestures *near* the object. For coarse gestures, it could be possible to rely on the hand capabilities of the Kinect API.

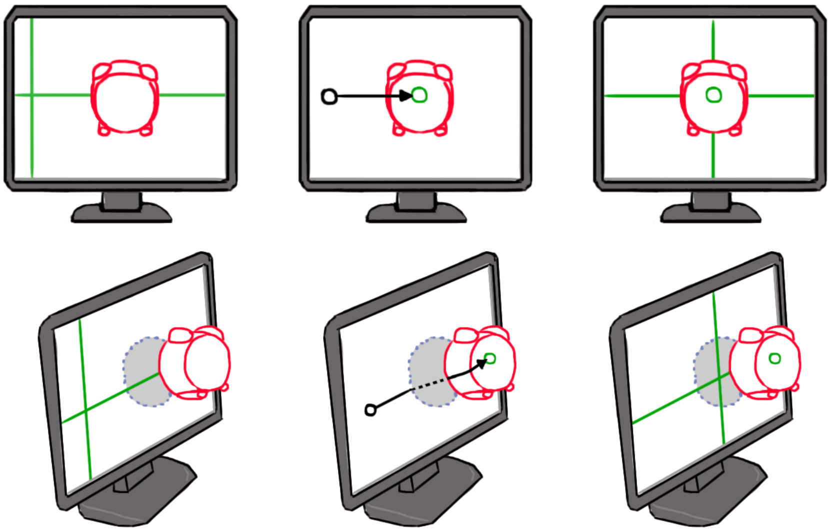


Figure 35: (Top Row) From the point of view of the user, the cursor behaves as if the object was part of a 3D viewport. (Bottom row) Side view showing the actual behavior of the mouse cursor, "jumping" from the screen onto the object when being occluded by the object from the user's viewpoint.

4.3 INTERACTION

In our hybrid workspace, the interaction can either take place on the screen, on the augmented object, or on both display supports at the same time. In the following sections, we explore the interaction space by describing techniques that we developed for each of these categories (Table 2).

4.3.1 *Screen*

Because our objective was to conceive a system that benefits from the advantages of standard desktops, all the usual techniques designed for such environments can directly be used.

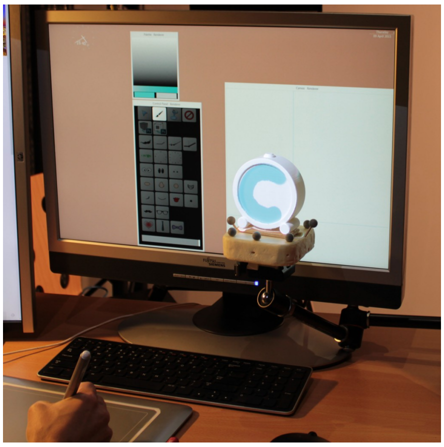
4.3.1.1 *Widgets*

We have designed a custom application based on such standard widgets for modifying the appearance of an augmented object (See Figure 36). For example, selecting the background color of an augmented object can be done directly by way of a color palette. This application served as a basis for the evaluation of the system that we present later in this chapter.

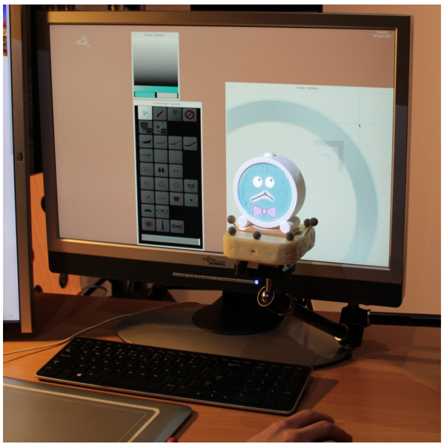
		MODALITY		
		Mouse and Keyboard	**Hybrid**	**Touch and Gesture**
LOCATION	**Screen**	Widgets, native applications, programming	–	Touch screen based interfaces*
	Hybrid	Drag and drop, hybrid widgets	Pick and drop, object annotations, data visualization	Gestural control of virtual version
	Object	Pointing on objects	Bimanual interaction	Navigation, tangible design

* Out of the scope

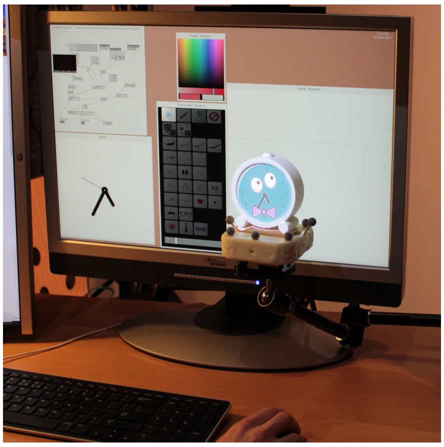
Table 2: Interaction space around Tangible Viewports.



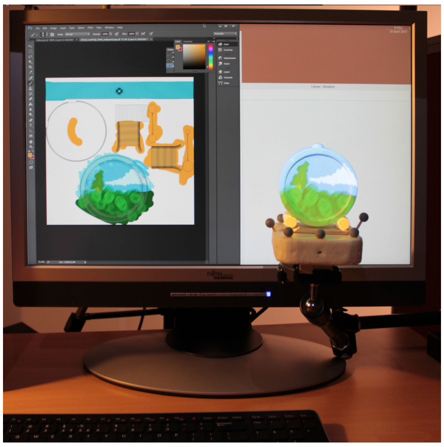
(a) Direct painting on the object using the mouse cursor



(b) Virtual elements interaction



(c) Creative coding



(d) UV painting using Photoshop

Figure 36: Different features to modify the appearance and behavior of the physical object.

4.3.1.2 *Native Applications*

It is also possible to use native professional applications. As an example, we linked the output of Adobe Photoshop, a software that is ubiquitously used in the design and artistic industries, to our system. Hence, we leverage the skills that professionals already acquired with these tools. The most straightforward use is UV painting (Figure 36d) which consists of editing the texture of a 3D model. It is a task that can be done either in a 2D painting environments using a UV layout or directly on a 3D view of the object. Both can be achieved using Photoshop. We retrieve the texture that is being painted in real-time and update the augmented object accordingly. Every time an operation is performed on the design, the physical object's appearance also gets updated. This can be especially useful in object design, where the final result is not a 3D render but an actual object.

4.3.1.3 *Programming*

In addition to the connection of existing tools, we also included Creative Coding capabilities. In practice, creative coding is often comprised of programming toolkits, such as the one discussed in Section 2.3.1, that are focused on visual results and short feedback loops. For these reasons, they are often used for prototyping. Taken in combination with the future envisioned by the OUIs community, where everyday physical objects are wrapped in thin and flexible screens (Section 1.3.4), we could imagine users being able to quickly tinker with the augmented content of their objects. Such content could be comprised of visual appearances, but also *behaviors*. Reality Editor [54, 55] and the derived open source toolkit Open Hybrid [1] are good examples of such programming frameworks. Holman and Benko [58] also combined tangible objects with programming in an installation that enabled the prototyping of tangible interfaces. However, more complex programming is an activity that is almost exclusively conveyed on standard computers. It is possible, then, to create a program and visualize its execution in real-time on a tangible object. As an example, we programmed a short example where the appearance of a clock evolves with the time. The results of this program can be visualized directly on an augmented physical clock (Figure 36c).

4.3.2 *Physical Object*

This section presents the interaction techniques we have implemented to support the use of physical objects: direct interaction, pointing on objects using the tangible viewport window and bimanual interaction.

1 http://www.openhybrid.org/

4.3.2.1 *Direct Interaction*

Working with physical objects has the benefit of enabling manipulation directly with the hands. No 2D to 3D mapping operations are required to create a desired point of view as is required in 3D desktop applications. Also, since the augmentation occurs on the surface of the object, changing the viewpoint can simply be achieved my moving the head. The user can thus observe the object in a natural way, which radically differs from what he or she is used to do with a virtual version of models displayed on flat screens. Also, direct touch can be used whenever precision or specific tools are not required. For example, when creating interactive objects, one can use interactors or trigger behaviors directly, similar to [1].

4.3.2.2 *Pointing on Objects*

In addition to direct manipulation of the tangible objects, our system supports cursor-based indirect interaction for completing interaction tasks onto the physical objects. These tasks can be pointing, drawing, selecting or moving virtual elements. Compared to an approach where the user would interact directly on the physical object, indirect interaction offers several complementary advantages. It does not require specific input devices, it is fast and accurate, and it integrates within the desktop workflow.

4.3.2.3 *Bimanual Interaction*

Handling the physical object and using the mouse can be achieved at the same time following a bimanual interaction approach [8]. The hand holding the object plays the role of the reference frame and assists the dominant hand which is dedicated to fine mouse movements. This approach leverages the precision and stability of 2D pointing and the easiness of 6 degrees of freedom manipulations of 3D objects.

4.3.3 *Hybrid Screen/Object*

Both the physical objects and the screen are part of the same working space. Consequently, it is possible to directly link operations on the screen with actions on the physical objects. The converse is also true. We present application examples that use both the object and the screen simultaneously.

4.3.3.1 *Drag and Drop*

Since the viewport creates a seamless continuum between the screen and the object, drag and drop operations can be used with the mouse cursor. It is interesting to note that this operation would not be possible using touch, and would have to be replaced by the pick and drop technique.

4.3.3.2 *Hybrid Widgets*

The standard approach for applying transformations (e.g. scaling and rotation) to visual elements displayed on a screen is to use widgets centered on these elements. The problem with standard SAR setups is that, although technically possible [16], it is very difficult to create the illusion of floating visual elements around the object when no material can support the display. We designed hybrid widgets that are operated on screen. From the user's point of view, they appear centered on the currently selected element. We reproject the position of the selected element on the screen based on the user's viewpoint and we place 2D widgets centered on this location. When moving the physical object, the position of the widgets is updated accordingly on the screen. These transformation widgets, which allow the rotation and scaling of the selected element are illustrated in Figure 36b. They are relatively big and they do not overlap the user's view of the physical objects. This design choice has been made to avoid problems of eye accommodation between the depth of the object and the depth of the screen. Hence, after selecting an object to modify, users can quickly grab and manipulate the widgets, without eye fatigue.

4.3.3.3 *Object Annotation*

Another opportunity offered by the choice of positioning the screen behind the tangible object is the display and the entry of text. Indeed, these operations may be difficult to complete in many traditional SAR setups. In our case, it is easy to annotate a physical object by selecting an anchor point (either with the mouse or direct touch) and typing a related note being displayed on the screen, with the keyboard – see Figure 37. Picking a point on the object creates a text box positioned in an empty zone of the screen which is linked to the projected position on the screen of the anchor point. Inversely, one can select a note on the screen, and see the corresponding area directly on the physical object.

4.3.3.4 *Physical Data Visualization*

Beyond annotations that can benefit to many fields (e.g. inspection of manufactured objects), we have explored the use of a hybrid approach in the scope of data visualization. Data visualization (and especially 3D data visualization) has been shown to gain from a physical representation [68]. Using the tangible viewport window, it is possible to add interactivity to physical visualization. In particular, to query more information on some aspects of the visualization, one can just point at the area of interest to see related data on the screen,

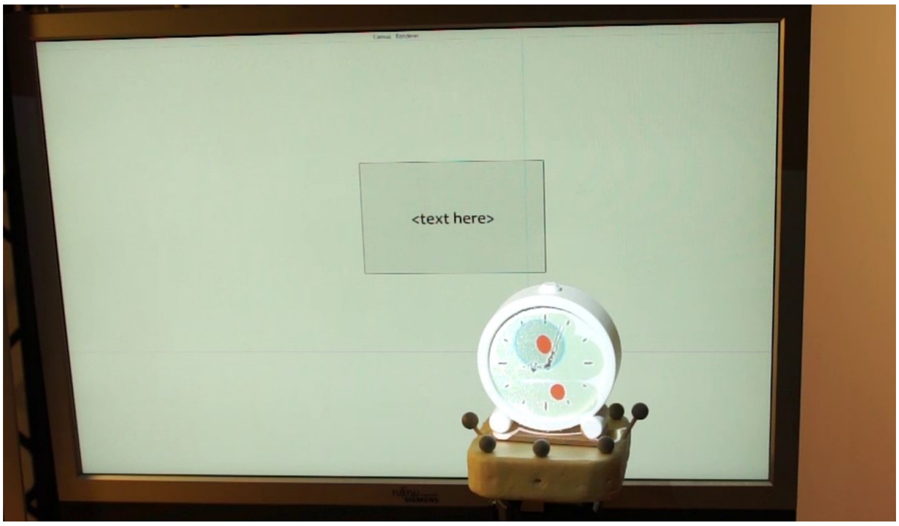


Figure 37: Textual annotations on a physical object leveraging the screen and standard inputs for pointing and text entry.

or she or he can select an entry on the screen to see the corresponding elements on the physical visualization (see Figure 38a).

4.3.3.5 *Synchronized Views*

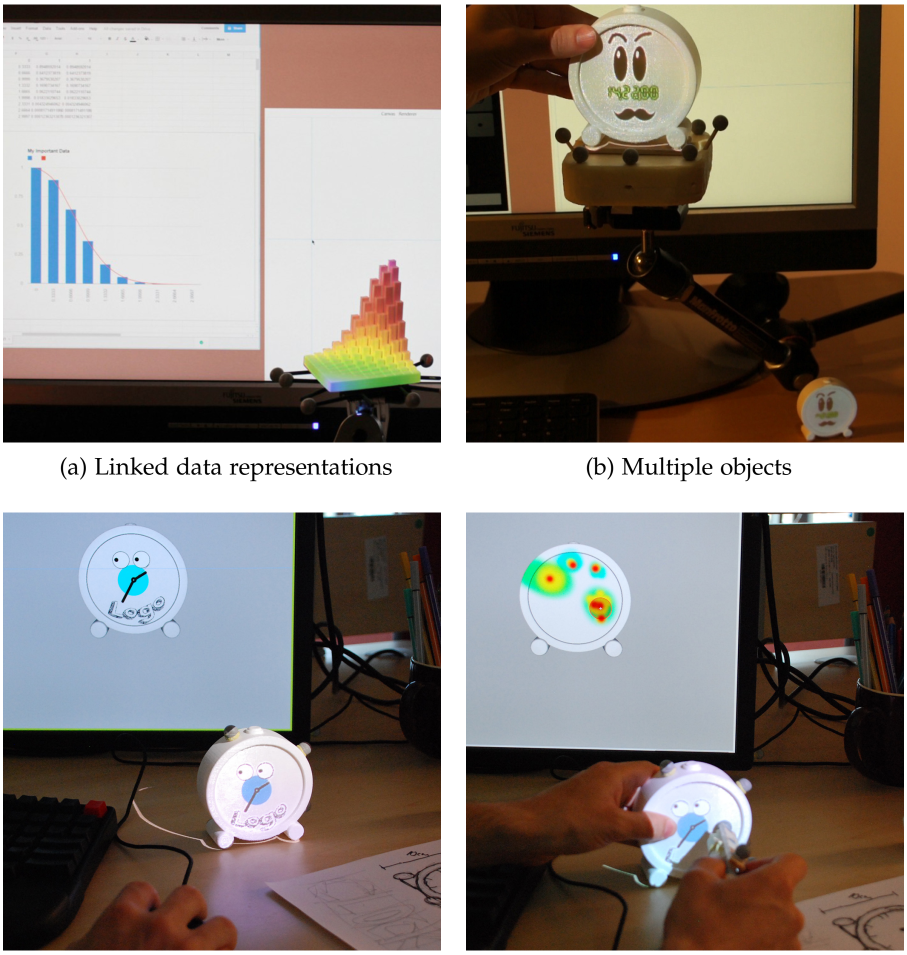
We also explored the synchronization between a virtual version of an object displayed on the screen and a physical one. When the tangible object is not in front of the screen, the tangible viewport window displays a virtual version of the augmented object (Figure 38a). Modifying the virtual version updates the tangible version in real-time.

Being able to have two representations of an augmented object, one on screen and one physical opens possibilities, namely for collaboration. For example, it would be possible to expose the view of a user handling the physical object or providing advanced visualizations such as a heat map of touched areas (Figure 38d). Also, multiple users can have their own duplicated augmented object (Figure 38b). These users can be working either locally or remotely.

The synchronization between real and virtual can be paused, for example using a gesture (e.g. pulling the object rapidly away from the window), to compare multiple versions. Bringing back the physical object in front of the window merges the two versions on the physical canvas.

4.4 ILLUSTRATIVE SCENARIO

To illustrate the use of Tangible Viewports, we describe here an object design scenario where an artist is experimenting with new visual design ideas for a product.



(a) Linked data representations (b) Multiple objects

(c) Synchronized views (d) Manipulation statistics

Figure 38: Different synchronization modes between virtual on-screen and tangible versions.

She can start by sketching first ideas on paper, and then use a modeling tool to create a 3D sketch. Equipped with a 3D printer, she can print one (or many) physical objects to have in front of her. She can first digitally paint directly on the object using the mouse cursor. Then, she can use a digital painting application such as Adobe Photoshop or a vector graphics editor like Inkscape to draft a logo on her computer. Then, using the mouse, seamlessly drag the logo from the editor directly to the physical prototype she just printed. The prototype can be physically manipulated to review the appearance. Modification to the visual design on the desktop computer will be automatically reflected in real-time onto the object. She can scale and rotate the logo directly on the physical object to see directly the impact of her modifications. This way, the feedback loop between the design activities (which require specialized software) and the validation of the effect it has in physical form can be greatly reduced. If required, new versions of physical objects can iteratively be 3D-printed, as we currently do with 2D printers when working on 2D documents. By making the interaction with the physical objects coherent with the traditional way of manipulating 3D information on a desktop computer, it is possible to leverage the experience of users with their professional tools, while at the same time adding the richness of tangibility and physical visualization.

For this illustrative scenario, it is worth mentioning the work of Saakes et al. [136], consisting in a physical material that can retain its appearance after being exposed to visible light. For an object design application, this has the potential to make the use of SAR a serious contender to generate visual designs that are not limited to dark environments and dependent on projectors being located in the room.

In our scenario, we can also imagine one or several collaborators participating to the design choices. These collaborators can directly observe and manipulate the augmented object, and ask the main designer to update the design in real time. This kind of social collaboration is harder to obtain with traditional design tools.

4.5 User feedback and discussion

We conducted an exploratory study where we asked participants to manipulate a preliminary version of the system, as well as a non-tangible version of the tool. The objective of this study was to assess how physical objects integrate within a standard screen space. We have designed a simple custom creation tool (see Figure 36a and 36b) for this purpose. Fourteen participants (9 males, 5 females, mean age 25.6 ($SD = 3.7$)) took part in this study. Half of the participants started the experiment with Tangible Viewport, then they moved to the non-tangible one, and half did the opposite. In both cases, participants were introduced to the main features of the tool, and the experimenter explained what was expected from them. Participants were asked to create a personal visual design of a clock. The only difference between the two versions of the tool is that, in the tangible version, the results of the creation was directly displayed on a 3D printed clock, whereas the virtual representation of the model was used in the standard viewport version. For changing the view on the object, subjects had to either manipulate the object and/or move the head "naturally" with Tangible Viewport, whereas they were using a trackball metaphor operated with the middle mouse button in the standard viewport version, as commonly done in standard desktop 3D tools.

Participants were asked to follow a tutorial for customizing their clock (Figure 39), which included: 1) choosing a background color and painting the front face, 2) adding virtual elements and resize/rotate them, and 3) making a drawing on the side and back of the object. This scenario was designed to ensure that the main features of the tool were used under different conditions. For example, Step 3 tests the ability of the participants to draw freely on curved surfaces.

After the experiment, participants were asked to answer two questionnaires: the User Experience Questionnaire [85] and a custom questionnaire aiming at obtaining user feedback about the usability of the tested systems (5 points Likert scale) and their preferences between the two. Both questionnaires showed no significant difference between the two versions of the system. Participants were also invited to leave comments and feedback about what they liked and disliked about each version of the system. Overall, the majority of the participants preferred manipulating the tangible version (12 out of 14) and were more satisfied with the final result (11/14). No participant mentioned difficulties moving from the screen to the object. These results seem to indicate that the tangible viewport metaphor works well, and it is comfortable to use.

Regarding the comments, among the most appreciated features spontaneously cited by the participants is the ability to work with a real object (9/14) and to have a physical view on the final product (6/14). For example, P1 liked that *"you can see the real object with the elements you draw. That way, you can observe the final product before it is produced"*. P9 mentioned that *"The creation feels much less virtual"* and that *"going from the screen to the object is fun"*. A few participants also insisted that they liked to be able to manipulate the object with their hands (5/14), while others found it uncomfortable (5/14): the tangible objects were attached to a tracked magnetic base that could then be connected to an articulated arm; this was later corrected by removing the base, and adding the markers directly on the object so that the objects could be handled directly. Complaints were made (5/14) regarding the fact that the editor lacked important features such as "undo" and "magic wand". This highlights that the interaction be-

(a) (b) (c) (d)

Figure 39: Examples of participants' creations using Tangible Viewports.

tween the screen and the object was working well enough that the main focus was about the painting features of the application.

Regarding the technical solution, several participants (6/14) mentioned that the augmentation calibration was not precise enough. This could be improved by using more advanced calibration solutions such as the one used by Jones et al. [71]. They also explicitly mentioned some delays and robustness issues on the head tracking (5/14). The second iteration of the system corrected these issues by replacing face tracking by skeleton based head tracking and better Kinect positioning. Regarding the cursor, some participants (4/14) did not like the fact that changing the head position was moving the cursor on the object, a side effect of using perspective cursor. This issue could be addressed with a system that would prevent the cursor on the object to move when the head position of the user changes and instead correct the on-screen cursor's position when it reaches the edge of the object's silhouette, from the user's point of view. Such alternatives will be studied in the future.

4.6 SUMMARY

In this chapter, we have introduced Tangible Viewports and we have described an effective implementation of this concept. A preliminary study showed that the overall usability of this system is good. It is important to note that Tangible Viewports do not aim at replacing existing systems. Indeed, we have shown that, from a technical and user point of view, the seamless integration of physical and virtual tools is not just feasible but can enrich both.

One of the current limitations of a tangible approach is the rigidity of the physical elements, which cannot (yet) be reshaped in real time. Our vision is that 3D printing will become as efficient as 2D printing in a near future[2]. Hence, one will be able to use the flexibility of virtual elements to explore variations of geometries, and use physical

2 A good example of this progress is the technology proposed by the company Carbon3D (http://carbon3d.com/) which advertise a printing speed increase ranging from 25 to 100 times compared to traditional material deposit methods.

elements as soon as he or she will require a perception of the created shape that goes beyond a simple rendering on a 2D screen.

In the future, we would like to assess more precisely how Tangible Viewports may leverage creativity in professional uses. This will require dedicated user studies with targeted users for investigating in more depth how interactive physicality impacts performance.

One of the current limitation of Tangible Viewports is the fact that interaction with physical objects is only possible when they are located *in front* of the screen. It would be important to investigate how to provide more spatial freedom to users: being able to interact with the physical object in any given location. A potentially good starting point for this investigation would be the combination of CurSAR with the system described in this chapter. Indeed, during our experimentation with the system, we realized that the ability to slide the cursor off the screen onto a physical object laid in front of it changes the perception of the screen itself. Instead of being considered as a self-contained space, the screen suddenly becomes a *spatial element*. Combining a perspective-based cursor and the technique used in our system could further extend the interaction space of traditional desktop environments.

Finally, Tangible Viewports leverage the desktop computer capabilities as part of a set of tools to interact with physical matter. In the future, we plan to go further, by merging the desktop on the workbench itself.

Part III

INTROSPECTION

Getting to know oneself is in itself a lifelong journey. The more time one person spends reflecting on one's own action, behavior and thoughts – practicing introspection –, the more power he gets to reshape himself as the human being he would like to be instead of suffering his own internal reflexes. Technology has the power to make us more efficient to accomplish tasks, but it can also help us better know ourselves by reflecting on what is happening inside our own bodies and minds.

5 TEEGI: TANGIBLE EEG INTERFACE

Figure 40: Teegi (Tangible EEG Interface) is a friendly interactive character that users can manipulate to observe and analyze their own brain activity in real-time.

In the last two chapters, we have explored different ways to interact with augmented content hosted on physical objects. AR is especially useful to add information to real-world objects and environments that would otherwise be unavailable to us. Its combination with tangible elements, and by using spatial displays, reinforce even more this link to reality. One interesting aspect of our lives that is both personal but at the same time mostly hidden to us is the internal processes of our bodies and minds. In this chapter, we will more specifically focus on letting users explore what happens in their brain in real-time, using a combination of Electroencephalography – EEG – and tangible interfaces. Instead of directly augmenting the user's body, we create a proxy persona – a tangible avatar – which acts like a "tangible brain mirror".

5.1 CONTEXT

EEG measures the brain activity of participants under the form of electrical currents, through the use of a set of electrodes connected to amplifiers and placed on the scalp [112]. This technology is widely used in medicine for diagnostic purposes. It is also increasingly explored in the field of Brain-Computer Interfaces (BCI), the goal of which is to enable a user to send input commands to interactive systems without any physical motor activities, by using brain activity alone [187]. BCI is an emerging research area in Human-Computer Interaction (HCI) that offers new opportunities for interaction, beyond standard input devices [152]. These emerging technologies are becoming increasingly more popular. It feeds into fears and dreams in the general public where many fantasies are linked to a misunderstanding of the strengths and weaknesses of such new technologies. *No, it is not possible to read thoughts!* But what can be done exactly? Our motivation is to provide a tool that allows one to better learn how EEG works, and to better understand the kinds of brain activity that can be detected in EEG signals. Beyond the knowledge of the brain that a user can acquire, we believe that a dedicated tool may help demystify BCI, and consequently, it may favor the development of such a promising field.

We followed a multidisciplinary approach, combining HCI (Spatial Augmented Reality, Tangible User Interfaces), Neurotechnologies (EEG, brain signal processing) and Psychology/Human sciences (Human Learning and Representations, Scientific Mediation) to design an interactive multimedia system that enables novice users to get to know more about something as complex as EEG signals and the brain, in an easy, engaging and informative way. Our final goal is to enhance learning efficiency and knowledge acquisition by letting users actively and individually manipulate and investigate the concept to be learned [170], i.e. EEG signals.

This gave birth to Teegi – Tangible EEG Interface –, a physical character that users can manipulate in a natural way to observe and analyze their own brain activity projected in real-time on the character's head (see Figure 40). Beyond the technical description of Teegi, this chapter depicts an exploratory study we conducted, which provides an experimental basis for discussions and future works. To our knowledge, Teegi is the first system to make EEG signals and brain activity easily accessible, interactive and understandable. This work is based on theoretical foundations, technical developments, and preliminary investigations.

5.2 NEUROIMAGING AND EEG

EEG signals are small electrical currents (in the μV range) that can be measured on the surface of the scalp [112]. They reflect the synchronous activity of millions of neurons from the brain cortex (i.e., the outer layer of the brain). Compared to alternative neuroimaging techniques, such as MagnetoEncephaloGraphy or functional Magnetic Resonance Imaging, EEG is simultaneously cheap, portable and provides good time resolution. Because of these advantages, EEG

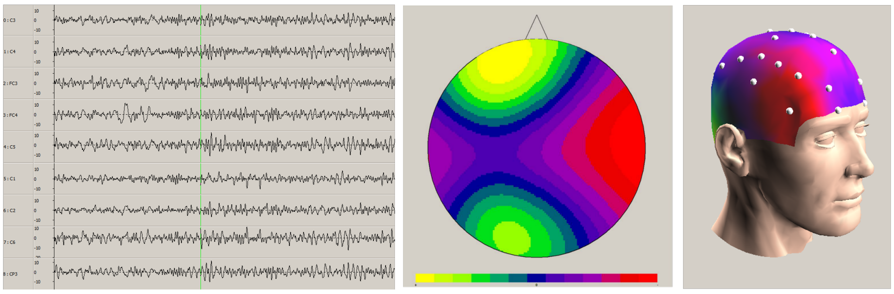


Figure 41: (Left) A trace of EEG signals collected from multiple sensors. 2D (center) and 3D (right) topographic maps. (Screenshots from OpenViBE [131]). The first two views are traditionally used by experts.

has been used for many years in medicine, e.g., for the diagnosis of sleep disorders or epilepsy [112]. More recently, with the advance of computer processing performance, it became possible to measure and analyze in real-time the content of these EEG signals. This paved the way for the rise of BCI which uses real-time analysis and decoding of EEG signals in order to identify the mental state of the user and translate it into a command for an application [187].

The currently available tools used to visualize and analyze such signals are tailored for experts with a deep understanding of the brain, EEG principles and EEG signal processing [112]. Figure 41 (left and center) shows some typical visualizations of EEG signals used by experts, i.e., EEG signal traces and a 2D topographic map. More complex visualizations have been proposed, such as 3D topographic maps (Figure 41, right), but they require many mouse inputs to be observed from all angles, which make them inconvenient to use in practice.

Although EEG visualizations are intended for experts only, the general public is often compelled by how the brain works and how its activity is measured. Anyone wondering about brain injuries, epilepsy, sleep or learning disorders, aging, etc. may want to seek further knowledge about how the brain works. Currently, the public is increasingly exposed to neurotechnologies due to the availability of consumer-grade EEG devices, such as the Emotiv EPOC or the Neurosky MindWave. Consequently, it has become necessary to design tools and user interfaces which will allow the general public to visualize, understand and interact with EEG signals. For instance, Mullen et al. proposed a software solution to process EEG signals collected using wearable EEG devices and visualize them in 3D [109]. This software enables the user to estimate brain activity sources and connectivity but is still mainly designed for brain signal and neuroscience experts, and not so public-friendly. Another recent work, more suited to lay persons, is the "Portable Brain Scanner" [147]. This system makes use of a consumer-grade EEG device (the Emotiv EPOC) and a smartphone to provide a cheap and portable solution enabling anyone to visualize the sources of their brain activity on their smartphone in 3D. Another more attractive work, which is the most closely related to our Teegi system, is the "Mind-Mirror" system [103]. This system combines Augmented Reality (AR), 3D Visualization, and EEG to enable users to visualize their own brain activity in real-time superimposed to their own head, thanks to a semi-transparent mirror-based AR setup.

This short review of the existing literature about making EEG accessible to the general public revealed that this is still a vastly unexplored area. Moreover, these solutions do not take into account any representation that the general public may have regarding the brain and EEG signals – many lay people do not even know what EEG signals are – in order to provide suitable visualizations and interaction devices to better understand these concepts. Some rare studies have indicated that misconceptions about brain functions prevail in general public [31, 53, 143]. These works stress the importance of popular scientific communication and indicate that communication efforts should be focused on increasing public awareness. It is important to note that the existing works mentioned above are mostly centered on visualization, with little or no interaction possibilities to manipulate and understand the EEG signals in real-time and in a friendly way. This further deters the general public from understanding brain activity [171]. Therefore, with the aim to enhance general public awareness, our work associates technical innovation and user-centered design.

5.3 INTRODUCING TEEGI

This section presents Teegi as well as the founding principles onto which the system builds. Moreover, we describe some advanced features created with the goal of increasing the expressiveness and control of Teegi, while keeping the tangible and accessible purpose of the system in mind.

5.3.1 *Founding Principles*

Design choices were made according to pedagogical principles. It has long been recognized that learner-centered education is much more effective than transmission-based education, even in informal situations [179]. According to the constructivist paradigm, people create unique personal meanings by reflecting on interactive learning experiences. Therefore, people/learners should investigate and manipulate in order to become conscious of complex phenomena, change their misconceptions and construct scientific knowledge [170]. In association, meaningful models play an important role in this type of learning processes [37]. This motivated the design of an anthropomorphic interface that can be freely manipulated

Our user-centered interactive media uses Spatial Augmented Reality (SAR) and tangible interaction. SAR, as introduced in chapter 2, is a technical mean to mesh the real and digital worlds together in a visually coherent experience. Along with TUIs – covered in Section 1.3.3–, this combination of approaches is particularly well suited for mediation purposes. One of the reason is their strong situatedness: the interaction takes place in a real-world environment that often hides most of the technological aspects to expose physical interaction components only. This tends to be more inviting to a general public compared to mouse-screen based interfaces [62].

There are examples of systems that use either tangible or AR principles to interact or review physiological data. Hinckley et al. [57] designed a system which used tangible props to do neurosurgical planning. A small tangible head was used in conjunction with a plastic plane to select the cutting planes to be visualized on a screen. BodyViz is a wearable system with organs represented as textile elements that are animated through embedded LEDs and screens to illustrate the inner working of bodily functions such as digestion to kids [116]. SWARM is also a modular wearable scarf that reacts to the wearer's emotional states by changing its temperature, emits vibration or sounds [182]. In a more sport-oriented context, Walmink et al. [173] proposed a bicycle helmet equipped with LEDs to display heart-rate information. Also mentioned above, the "Mind-Mirror" [103] is the work closest to Teegi. However, with Teegi, the data is not co-localized with the data source. It provides flexibility and easier visualization as the users can change viewpoints by tangible interactions instead of rotating their head while keeping their eyes on the mirror. This "out-of-body" visualization also enables collaboration where multiple users can explore the data.

5.3.2 *General Description*

Teegi is a tangible interface that enables users to visualize and analyze a representation of their own brain activity recorded via an EEG system in real-time and displayed on a physical character. After some processing of the raw signals, a dedicated visualization is projected directly on top of the character. This character is tracked, which allows us to co-locate the projection with the character's head, at any time. Hence, the user can easily visualize a realistic modeling of the EEG signals in any part of the scalp by manipulating the character, while maintaining a good spatial topology of the observed data. Teegi was purposely given a child-like appearance, as well as animated eyes (also projected) that blink at the same time as the users do, in order to breathe life to the character and enhance attractiveness. Indeed, blinking can be easily detected in electrodes neighboring the eyes.

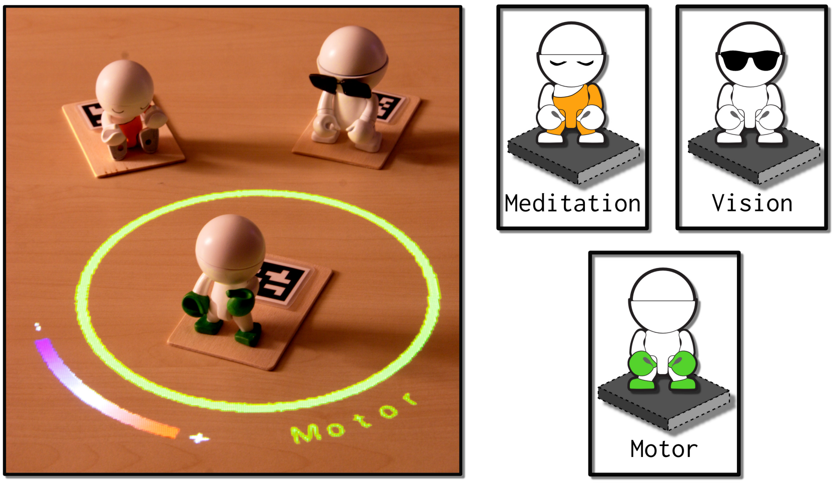


Figure 42: Three mini-Teegis can be used to apply high-level EEG filters to highlight brain processes associated to *Motor*, *Vision* and *Meditation* activities. To do so, the user simply needs to move the desired mini-Teegi into a specific zone projected on the table (green circle).

Three different filters can be applied to the raw data (see the technical section for details) enabling users to investigate influences of motor motions, visual activities or meditation, on their brain activity in real-time. To remain consistent with the tangible philosophy of this project, we decided to control the filters by way of small tangible characters (mini-Teegis) that can be moved on a "filter area", which is highlighted on the table by a projected halo (see Figure 42). For example, if a user wants to apply a filter that will allow her to better see what happens when moving her hand, she just needs to take the dedicated mini-Teegi, i.e. the one with the colored hands, and to move it to the filter area. Then, by moving her right hand, she should see changes in EEG amplitude on the left hemisphere of Teegi's head, as illustrated in Figure 43. The manipulation of Teegi requires a motor activity. Therefore, when the motor filter is on, manipulating Teegi will obviously lead to observable changes in brain activity.

At the end of this chapter, we present an explorative study we conducted to obtain feedback about the main features of Teegi. However, Teegi is not limited to these first features. In the next section, we describe additional interaction metaphors we have explored, and that may benefit more advanced users. These advanced features were not evaluated during the study.

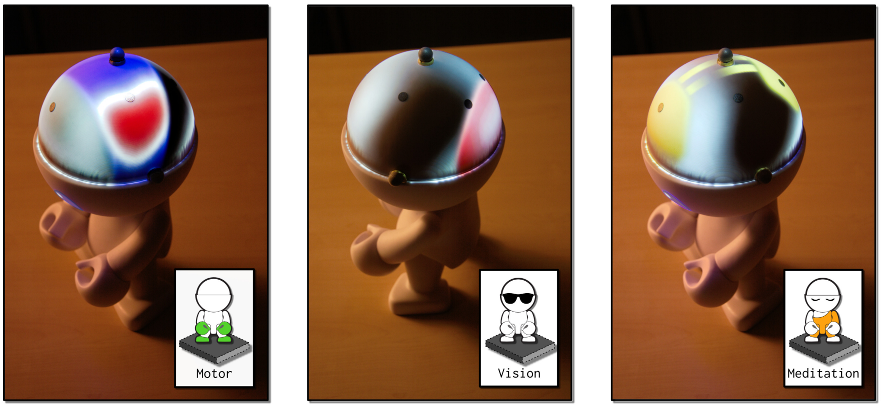


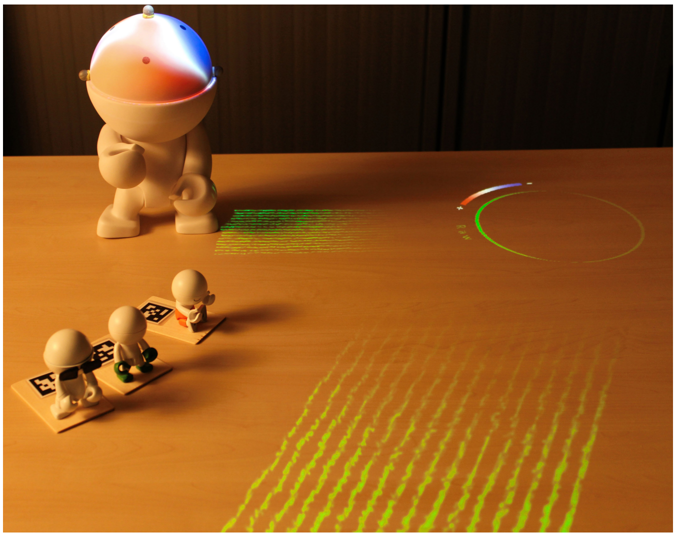
Figure 43: Examples of the displayed visualizations on Teegi for each of the provided filters. Once a filter is active, the brain area corresponding to the selected and processed activity is highlighted in colors while the remaining EEG signals are displayed in grayscale.

5.3.3 *Advanced Features*

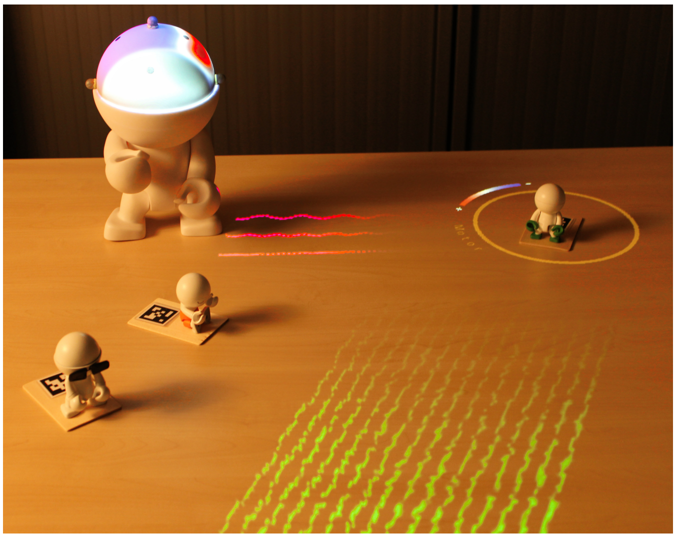
Visualizing the raw signal recorded on each electrode of the EEG is not very informative for the general public. However, this can be instructive for students who are learning EEG signal processing and analysis. In our approach, we can display on the table these raw data, as shown in Figure 44a. This creates a visual link between what is recorded with the EEG system, and the visualization that is provided on Teegi's head. This is possible because we know the rough position of the user and the exact position of Teegi. When applying a filter, as described in the previous section, the user can see the effect of his or her action on the signal (see Figure 44b). Compared to a standard approach where everything takes place on a screen, we believe that such a spatial and tangible approach might ease the understanding of the filters' effect.

Another dimension we explored is the use of tangible actions to control some parameters of the EEG signal processing. As an example, we have implemented a technique where the user can control the amplitude of the visualization color map by moving a tangible object on the table (Figure 45). This could be useful to reveal tiny fluctuations of EEG signals. With such interaction techniques, the whole analysis could be conducted without the use of a screen or a mouse, which remains consistent with the tangible philosophy of the project.

Finally, we developed a solution that highlights the relationship between EEG signals and localized cortical sources, that is where the signals come from *inside* the brain. Using sLORETA inverse modeling [120] and Brainstorm to compute the kernel matrix [150], we obtained a model of the cortex containing 2002 voxels linked to the 32 EEG electrodes we used. We can then project in real time the activity which arises from the outer regions of the cortex on an object



(a) Raw EEG readings



(b) Filtered signals

Figure 44: Exposing the signal filtering process. *(a)* The raw EEG readings are displayed going from the user to the filter area and then rerouted towards Teegi. *(b)* When a mini-Teegi (i.e. a filter) is active, the corresponding filtered signals are displayed between the filter area and Teegi instead.

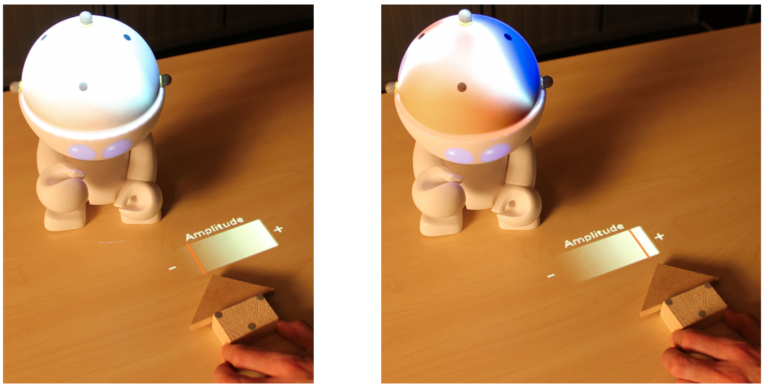


Figure 45: A moving tangible cursor is controlling the amplitude of the visualization color map.

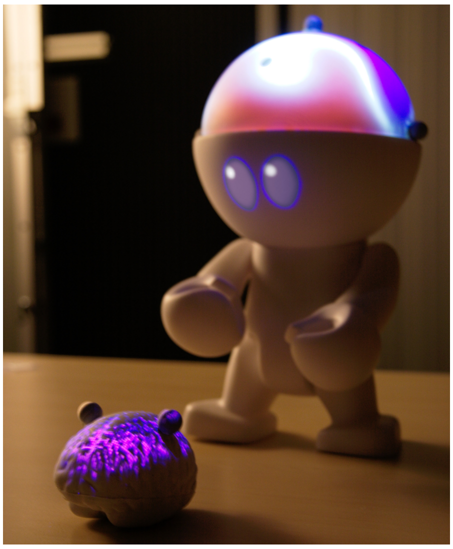


Figure 46: Using an inverse model, the cortical activity and EEG measures are presented together to users.

representing the brain, alongside with Teegi (Figure 46). Since both Teegi and the brain proxy are tracked, it becomes possible to manipulate two synchronized representations of the same brain activity (the source at the surface of the brain and the measures on the scalp). This opens the way to mediation activities that are more advanced all the while keeping the simplicity and ease of use brought forth by using SAR and tangible interaction.

5.4 TECHNICAL DESCRIPTION

In this section, we present how the Teegi system was built. We will cover the aspects related to EEG measurements and related signal processing as well as the visual augmentation using SAR.

5.4.1 *EEG*

We designed different EEG signal processing pipelines that each creates a specific visualization tailored to identify specific elements in the signal. The details of these pipelines are transparent to the user. Each pipeline corresponds to a mini-Teegi filter. In particular, we set up the following EEG signal processing pipelines:

WIDE-BAND EEG ACTIVITY: EEG signals were band-pass filtered between 3 Hz and 26 Hz, in order to filter DC drift and part of the artifacts (e.g., facial muscle activity [35]) that may pollute them. Their power is then computed before being displayed. This corresponds to unspecific brain signals, hence they were labeled as "raw" signals.

SENSORIMOTOR ACTIVITY: EEG signals were first band-pass filtered in the β band (16-24Hz), a brain rhythm highly involved in sensorimotor tasks [122]. Then, they were spatially filtered, i.e., the signals from several neighboring EEG sensors were combined in order to enhance the signal of interest. In particular, we used and displayed Laplacian spatial filters around electrodes C3, C4 and Cz. This enabled the users to visualize EEG activity changes due to movements of the left hand, right hand and feet. Indeed, it is known that the power of EEG signals in the β rhythm decreases in electrodes C3/Cz/C4 during right hand/feet/left hand movements respectively, and increases just after the end of this movement [122].

VISUAL ACTIVITY: EEG signals were band-pass filtered in the α band (8-12 Hz), then only electrodes P3, Pz, P4, PO3, POz, PO4, O1, Oz and O2 (located on the back of the head, above the neck) were selected and displayed. These electrodes are indeed located over the visual cortex of the brain, i.e., the brain area in charge of visual information processing. The amplitude of the α rhythm is actually known to increase while the user is closing his/her eyes and is thus not processing any visual information [112]. To ensure that the user could perceive this increase after he/she reopened his/her eyes, the visualization was delayed by 0.5s.

MEDITATION: on a more exploratory note, we used the synchronization between the signals from the anterior and posterior cortex (AFz/Pz), which was measured in a 7-28 Hz band with instantaneous phase locking value [83]. There are different outcomes (increase/decrease in synchronization) depending on meditation type. Mindfulness and body focus practices decrease the synchronization while transcendental practice increases it [90].

EEG signals were acquired with a 32-channels EEG device (made of two g.tec g.USBAmp EEG amplifiers). This professional-grade system ensured that our prototype had a good signal-to-noise ratio and accurate electrode location, avoiding unneeded uncertainties. Signals were processed in real-time using OpenViBE [131]. For pipelines 1 to 3, the displayed colors correspond to signal power strength; for pipeline 4 they correspond to the degree of synchronization.

5.4.2 *Spatial Augmented Reality*

In order to create an augmented character, we have designed a tabletop augmentation setup (see Figure 47). Teegi itself is a 25cm high Trexi DIY toy. The mini-Teegis are also 10cm high Trexis. The main program handling the whole installation was created with vvvv[1]. The primary projected content (Teegi augmentation and GUI display) is handled with a single wide lens projector ProjectionDesign F20SX of resolution 1024x768 located over the table in a top-down orientation. The tracking of Teegi is achieved with an OptiTrack V120:Trio. It runs at 120 FPS with an overall latency of 8.3ms and a precision of 0.8mm. The OptiTrack is located in the same configuration as the main projector and both devices are calibrated together manually. The tracking data is sent to vvvv using OptiTrack's NatNet protocol. Teegi's eyes are projected using a second projector (Vivitek Qumi Q2) that is located on the side of the table.

The filter selection is done using a Sony PSEye web camera pointed at the position of the program selection GUI. Each mini-teegi representing a filter has a fiducial marker attached to it. The library ARToolkitPlus [172] is used to detect which marker is currently selected.

The EEG signals are processed by the OpenViBE software [131] that also generates a grayscale texture of the scalp signals. This texture is then exported to a local shared samba folder which is then fetched and remapped to an appropriate color scale in vvvv before being mapped to Teegi's head. In addition, the raw EEG signals are sent to vvvv over VRPN for display purposes (see Figure 44).

5.5 EXPLORATORY STUDY AND DISCUSSION

We conducted an explorative study where participants had to manipulate Teegi following a given scenario. The objectives of this study were to i) evaluate the general usability of the interface and ii) obtain initial feedback about the relevance of the approach to help users understand EEG signals and the brain.

1 http://vvvv.org

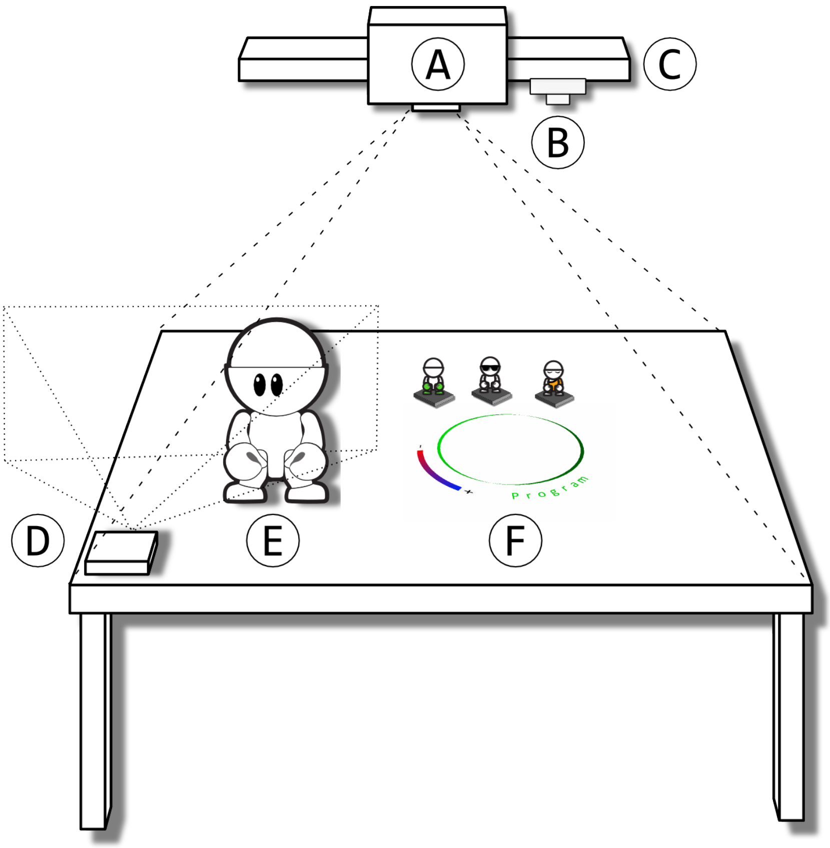


Figure 47: Diagram of the installation. (A) ProjectionDesign F20SX projector (B) Sony PSEye web camera (C) OptiTrack V120:Trio (D) Vivitek Qumi Q2 projector (E) Teegi (F) Program selection zone and mini-Teegis.

5.5.1 *Protocol*

Ten participants (6 males, 4 females, mean age 28.6 (SD=9.7)) took part in this study. Pre-tests confirmed they were rather naive on the subject. They manipulated the version of Teegi described in the General Description section (no advanced features). The general procedure was as follows:

1. Pre-tests: The participant answered a first questionnaire assessing his or her representation of the brain. The participant then filled in different forms to measure his or her previous knowledge; one form per studied brain process (motor, vision and meditation).

2. Setting-up: The experimenter positioned the EEG cap on the participant's head. In parallel, the participant, guided by the experimenter, was made aware of the four didactic "cards" explaining the different filters i.e., *Motor*, *Vision*, *Meditation* and *Raw*. Each card was comprised of an image of the mini-Teegi associated with the filter along with basic instructions to follow (e.g. the *Motor* card indicated to the participants to move their hands or feet while staying relaxed). There were also two cards describing the two types of visualization participants could face, *signal strength* and *synchronization*. Once the participant was equipped, a quick calibration phase occurred. While Teegi was still inactive, participants were asked to close their eyes for a few seconds, and to move their hands and feet in order to identify the baseline activity for visualization.

3. Personal Investigation: The participant was asked to freely manipulate Teegi as well as the filters to be able to answer the following questions:

   - What happens when you move your hands or feet?
   - What happens when you close your eyes?
   - What happens when you meditate?

   During the whole study the participant sat comfortably in a chair. To avoid the occurrence of muscle artifacts that may pollute the signals, the user was instructed to stay relaxed and to refrain from making strong head movements.

4. Post-tests: The participant answered the questions above on dedicated forms, the same that were given at the beginning of step 1. Finally, he or she filled in a user survey questionnaire based on a 7-point Likert scale.

The whole session lasted approximately 1.5 hours per participant, with 15 to 20 minutes of hands-on time with Teegi. Each session was video-recorded. Video segments were separately visualized and labeled with the corresponding behavior (i.e. tangible and visual interactions, emotional expressions, and investigation strategies) using The Observer XT® 11.5[2]. After the session, the experimenter had an informal talk with the participant. He corrected the answers, making sure the participant was not leaving with false knowledge, and explained in more detail some aspects of the system (e.g. relationship between visual filter and attentional states, the various effects of meditation, ...). This phase lasted from 30 min to 1 hour depending on the participant's curiosity.

5.5.2 *Results and Discussion*

To better understand the inherent strengths of Teegi towards learning, we assessed three main aspects of Teegi: its technical reliability, its relevance to ease understanding for non-experts, and the User eXperience (UX) it provides. This evaluation is based on:

1. The results of the questionnaire that are summarized in Figure 48;
2. the analysis of the video recordings;
3. the analysis of the forms the participants filled in to assess their pre and post-knowledge of the brain and EEG.

5.5.2.1 *Technical Reliability*

Participants unanimously reported that the whole system worked properly. The quality of the SAR display is valued by the participants. In particular, they reported that the resolution was appropriate, and they did not report problems of offset between the display and the physical character. Participants declared that they were not disturbed by occlusion problems. The mild temporal delay between their action and their consequences seems not to be an issue.

Manipulations of Teegi were numerous and frequent. Teegi was touched or moved on average 25% of the session's duration, twice per minute. These manipulations consisted mostly of rotations, and to a lesser extent of lifting Teegi to enhance visual perception. Two participants reported difficulties in grasping Teegi while the remaining 8 were comfortable with the form of the character. Video analyzes did not show difficulties for the manipulation of Teegi. Similarly, applying filters by manipulating the mini-Teegis seemed easy for the participants.

2 Noldus, Info Tech, Wageningen, The Netherlands

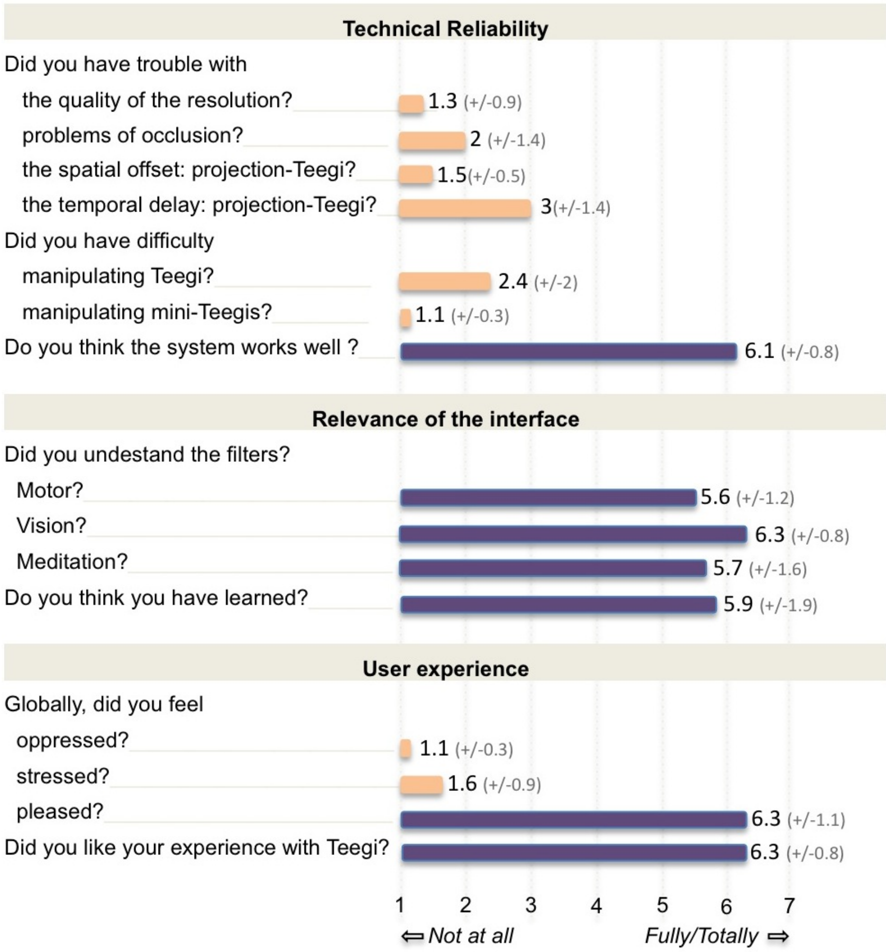


Figure 48: Results of the questionnaire (selected questions). Note that purple (resp. orange) bars indicate questions measuring Teegi's qualities (resp. limitations).

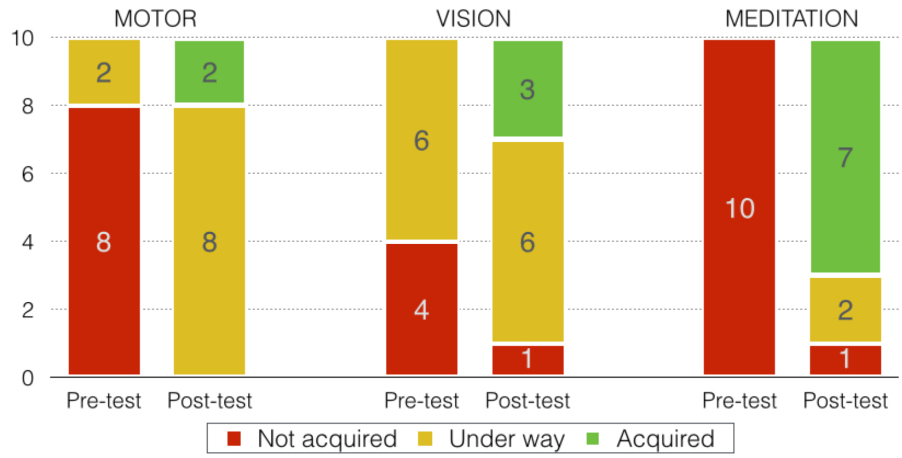


Figure 49: Marks obtained by the participants during the pre- and post-test assessments. See text for details.

5.5.2.2 *Relevance of the Interface to Ease Understanding*

The participants reported that they understood the visualization associated with the filters. Video analyses indicated that they systematically used all filters several times (3 times per session on average) for a similar duration (Raw filter : 30.4% (SD 13.3) of session duration; Motor filter: 26.0% (8.3); visual filter: 16.9% (5.5); meditation filter: 26.6% (8.7)). Interestingly, the visual activity filter seemed slightly easier to understand than the other filters. Moreover, video analyzes indicated that the participants did not have difficulty observing the signals on Teegi's head, as soon as they found the right location to observe. Overall, participants reported that they were able to use Teegi without any difficulties.

All participants completed the required tasks. They used instruction cards 5 times per session on average. They reported that they could focus on the tasks rather than on the mechanisms used to achieve them. This suggests that Teegi is a rather transparent interface. Regarding learning of brain processes and EEG, participants reported that they believed they had learned while doing the study. This was confirmed by the results of the pre- and post-test assessments (see Figure 49). These assessments focused on the recognition and the understanding of brain activation during Motor activities, Visual activities and Meditation. Understanding was marked as acquired if 1) the activated areas were correctly localized and 2) the explanations of the brain process were correct. It was marked as under way if only 1) or 2) was satisfied but not both; and as not acquired if neither 1) nor 2) were satisfied. The marks obtained by the participants improved after using Teegi. Overall, this suggests that Teegi offers many interesting features to ease learning and mediation.

All our results indicate that Teegi clearly promotes real-time tangible interactions, which contributes to enhancing awareness. Constructivism and inquiry-based science education principles indicate that in order to become conscious of complex phenomena and construct scientific knowledge, people/learners have to experiment by interacting with and physically manipulating the content [171]. This is particularly true for brain activity that is difficult to understand because it cannot be sensed [29], contrary to other activities (e.g. respiratory) that are perceived through sensory-motor mechanisms. Hence, brain activities need to be conceptualized, and the success of learning processes strongly depends on the interface. Teegi, which has been largely promoted by the participants, seems to fulfill this function.

5.5.2.3 *User eXperience*

The general experience with Teegi was rated as pleasant, attractive and stimulating, and participants did not feel stressed or oppressed. Overall, participants reported that they liked interacting with Teegi. The emotion expression analyzes confirmed those statements. They showed that on average participants expressed curiosity and questioning about Teegi feedback during almost 20% (20.1% SD=9.1) of the manipulation duration. Other emotion expressions observed for all participants were joy and pleasure (e.g. smile, laugh, joyful verbal expression). They occurred during almost 10% (9.8% SD= 6.7) of the interaction duration with Teegi. Surprise emotions were observed but less frequently. Interestingly, boredom, weariness expressions occurred rarely (only for 2 users) and only at the end of the manipulation time. We did not observe any occurrence of exasperation or irritation. These results suggest a high level of acceptance for Teegi. This is a fundamental requirement for a tool aiming at improving access to knowledge.

Behavior observations indicated that the majority of participants spoke with Teegi and used morphological zones specific to human interactions while manipulating it. For example, they held its hands and held it up by the waist as one would do with a child. Some users spoke in the first person when they observed changes on the character's scalp for example "so, when I move my hands, I light up on the sides"; many said aloud that Teegi was their own image, for example "so, Teegi is me!". This identification suggests that an activation of associations between the perceived character's personality and self-perception may have occurred [119]. It is known that identification can be associated with increasing loss of self-awareness and its temporary replacement with elements of the perceived character's personality [27]. Therefore a human shaped, child-like character, made lifelike by animated projected eyes, could enhance both empathy and implicit self-perception of one's own brain activity, as provided by our interactive media. The anthropomorphic appearance of Teegi could explain the motivation and positive UX reported by the users. All these hypotheses would be the aim of a more extensive UX study.

Regarding visual attention, the participants were apparently paying attention to Teegi most of the time (83.3%, SD 7.6). This supports the fact that Teegi mobilized user attention. It also indicates a cognitive user engagement. Personal investigations were permanent (only 1.9% of inactivity was measured during the session duration; SD=1.7). Behavior analyzes indicate that participants made predictions, hypotheses and tested them by conducting experiments. Numerous trial and error strategies were frequently used. This clearly indicates personal active control of the task and inquiry processes. Overall, Teegi stimulates investigations and encourages persistence in task completion.

5.6 SUMMARY

In this chapter, we presented Teegi, a tangible interface that makes EEG understandable to non-expert users. Our main contribution is the interface itself, which is built from both theoretical foundations, notably from human learning and scientific mediation and technical developments, including spatial augmented reality, tangible interaction and real-time neurotechnologies. We also demonstrated that this interface was well accepted by a first pool of users. In the future, we plan to make a more in-depth investigation into how well users are able to learn about EEG and brain activity with Teegi. To this end, we will conduct dedicated experiments with students and/or visitors in scientific museums. We would also like to precisely evaluate how Teegi benefits learning compared to standard approaches. For more advanced users, ad-hoc tangible filter creation could prove to be of great interest, adding flexibility to the overall system. It is also known that BCI requires the user to learn to control his/her own brain activity to input computer commands [187], which is a long and tedious task. We expect Teegi to be a motivating and informative way to support this training in the future.

Finally, while Teegi is focused on EEG visualization, it offers a good starting point and initial platform to widen its purpose. Indeed, in the next chapter, we will discuss how we applied our experience with Teegi to create an open platform that exposes a greater range of physiological states and metrics.

6 TOBE: TANGIBLE OUT-OF-BODY EXPERIENCE

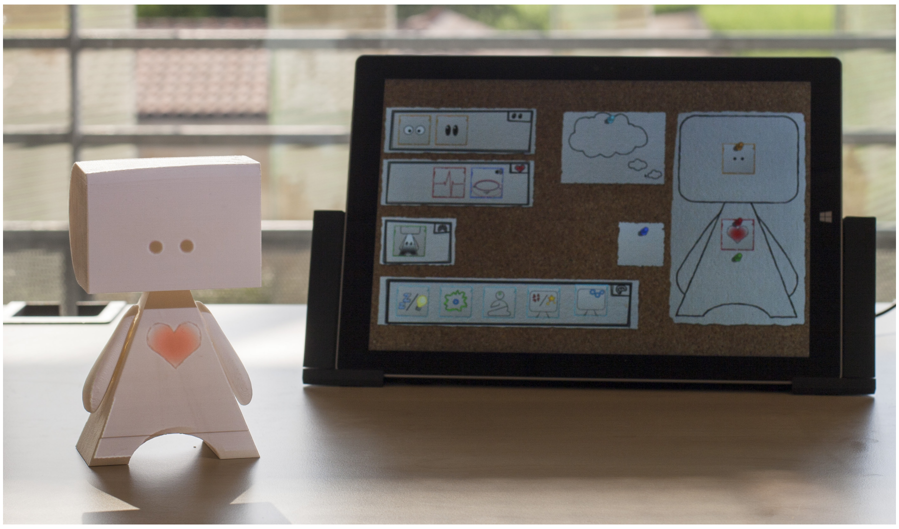


Figure 50: Tobe, the tangible avatar displaying real-time physiological readings along with the interface to control the different visualizations.

This chapter introduces Tobe, a toolkit for creating Tangible "Out-of-Body" Experiences. These tools are presented using one incarnation of the system based on a tangible anthropomorphic avatar and spatial augmented reality to display real-time physiological signals and mental states (Figure 50). The main objective of this project was to enable users to know more about themselves and others using a representation that they could shape themselves – that is, creating a tangible avatar of their *inner* selves.

6.1 CONTEXT

Wearable computational devices are more accessible and more popular than ever. These devices are personal and are embedded with physiological sensors. The most current use of these sensors are health monitoring, e.g. runners have access to their heart rates, traveled distances and their itinerary. Physiological sensors measure something deeply personal while at the same time measuring information that is not explicitly available to us. The "Quantified Self" (QS) [186] movement arose from the availability of such sensors [92] and people uses them to collect large amount of data about their lives in order to reflect on them. This trend takes place in an even larger movement of personal life improvements. We indeed live in an era of Do-It-Yourself, "life hacking" and personal development. While QS has been mostly about collecting physical health data, extensions to cognitive tasks has also been proposed [82]. Such data collection is popular exactly because people are not explicitly and objectively aware of how well their bodies and minds are doing. However, these applications are limited compared to what physiological sensors can offer. Nowadays even brain activity is within reach of consumers thanks to cheap alternatives to medical equipment, such as the Emotiv Epoc[1] or, closer to the Do-It-Yourself community, the OpenBCI board[2]. Indeed, physiological computing is mature enough to assess mental states [34, 39, 123, 192]. Therefore, it could be used as a mean to better know our own self and others.

On the one hand, physiological technologies are not exploited to their full potential and on the other hand, we have end-users that ignore what technology has to offer for their well-being. Some companies are pioneers, as for example Empatica and its Embrace smart watch[3], but such companies focus on health applications and, consequently, the targeted consumers are still a niche. Both a process that will raise public awareness and a collection of meaningful use cases are missing. Finally, when bodily activity and mental states are at stake – which are difficult to conceptualize and often difficult to perceive – the feedback given to users matters for them to comprehend at first sight what is being measured. How to represent the arousal state of someone? How would *you* represent cognitive workload? We found little examples besides pies and charts, which are not always obvious informants in data visualizations – e.g., [100].

To address these issues, we first conducted surveys and interviews to gain insight about physiological feedback. We then created Tobe (to be pronounced [`tobi`]), a Tangible Out-of-Body Experience shaped as a tangible avatar (Figure 50). This avatar lets users freely explore and represent their physiological signals, displayed on the avatar itself using spatial augmented reality. The overarching goal is to help one reflect on his physiological and mental states in *his or her* own way. The main activity would be for users to actively *build* from the ground up their own self-representation and then visualize physiological signals through it. As such, we designed a modular toolkit around Tobe that can be used to customize any part of the system. Tobe has been tested on two different occasions in a scientific museum (*Cap Sciences*[4], located in Bordeaux, France) to collect user feedback. A specialized version of the system was also built to give biofeedback to multiple users in a relaxation task. Beside these two implementations, we frame potential uses of the system, such as a biofeedback

1 https://emotiv.com/
2 http://www.openbci.com
3 https://www.empatica.com/
4 http://www.cap-sciences.net/

device for stroke rehabilitation or replaying inner states synchronized along with videos of cherished memories. The latter example could help create more cherishable versions of personal digital data [46].

One of our objectives with Tobe was to foster reflection about well-being and inner states. While the toolkit we propose can definitely fit well into the "Quantified Self" movement as a way to represent and collect data in a tangible way – as it was done in [76]– we wanted to go for a *Qualified* Self. Qualitative representations of information, especially when on a continuum, have the advantage of avoiding explicit metrics. Numbers have this property that they can easily be compared – that is how scores work. A score can be used to compare against our own previous performances or performances of others. This implicitly creates competition, a state more or less compatible with calm, reflection and well-being [20]. Moreover, introspection is an activity where one is focusing on his or her own states, trying to avoid comparing against others.

Previous works do not embrace such system as a whole and are limited either to low-level signals or to emotions. Wearables were used to mediate affect in [182] – a context similar to the "Social Skin" project [160] – or to teach children how the body works [116]. Tangible proxies were studied in [76], but the feedback was not dynamic. In [38] a tangible puppet was already used as a proxy for brain activity, but the settings concerned scientific outreach and the feedback focused only on preprocessed brain signals and not on higher level mental states (which Tobe does). Our toolkit pushes further the boundaries of the applications by giving access to physiological signals, high-level mental states, dynamic and customizable feedback.

Our contribution for this work consists of a toolkit enabling users to create an animated tangible representation of their inner states. The toolkit encompasses the whole workflow, including the physical avatar creation, sensors, signal processing, feedback and augmentation. It was tested through two use cases, in public settings and with a multiple users scenario.

6.2 REPRESENTING PHYSIOLOGICAL SIGNALS

Exposing physiological signals in a way that makes sense for the user is not trivial. Some types of signals might be more obvious to represent than others. For example, heart activity could be understood using a symbolic heart shape due to largely accepted cultural references. This question is, however, harder when talking about more abstract mental states such as workload. Nevertheless, even the dynamic representation of low-level physiological signals is still an open question at the moment [25]. We conducted two surveys to gain more insight about the knowledge and the representation people had of different types of signals and high-level mental states.

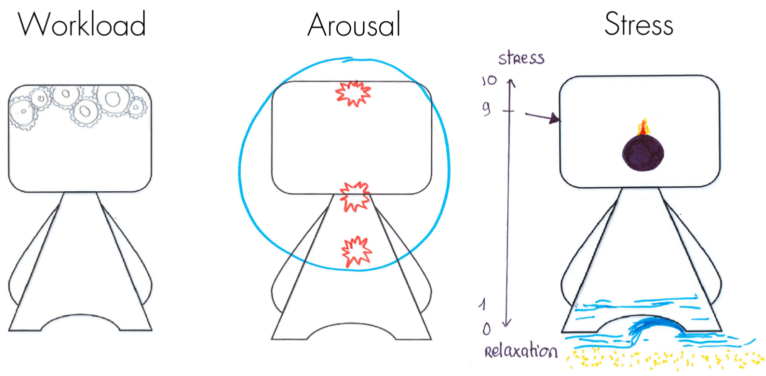


Figure 51: Sample of the drawings made by participants to represent various high-level metrics.

In the first survey, conducted online, we asked 36 persons about their knowledge of physiological signals in general. We inquired about the self-awareness of inner states on a 7-points Likert scale (1: no awareness, 7: perfectly aware). About "internal physiological activities", the average score was 3.5 (SD=1.4) and for "mental states", the average score was 4.9 (SD=1.3). The latter score indicates that the participants thought they *knew* their inner states – even though a whole literature demonstrates how difficult this is [113] and how easily we are deceived [161, 72]. Interestingly, we also observed that most of the participants reduced mental state and physiology to emotions only. Mentions of any cognitive processes such as vigilance and workload were very rare (7 out of 36). This lack of knowledge about the inner self and the different cognitive processes is an opportunity to raise awareness of the general public about the complexity of the mind. When inquired about possible uses of a Tobe system, very few respondents (6 out of 36) gave examples other than sports or health. This emphasis the fact that the general public is unaware of possibilities of technology for well-being.

The second survey specifically investigated how users would shape the feedback. We focused on visual cues because it was easier to express on paper, but note that other modalities could be explored, such as sound [102]. We asked 15 participants to express with drawings and text how they would represent various metrics (Figure 51). There was little resemblance between subjects for a given high-level metric and even low-level ones – breathing and heart activity – sprang different views. For example, some people drew a physiologically accurate heart instead of a simple sketch. Overall, there was a wide variety of sketches and people were very creative. This highlighted the absence of consensus on how we conceive and view our inner states. Therefore, people could benefit from being able to tailor a meaningful and personal feedback.

6.3 TOOLKIT

We created a tangible anthropomorphic avatar, named Tobe as a host for displaying real-time feedback. We chose this form factor because we found evidence in the literature that this combination of anthropomorphism and tangibility can foster social presence and likability [139, 62, 63]. This also reminds users and observers that the feedback is linked to an actual being; it helps to recognize Tobe as a persona and to bond with it, hence it facilitates engagement.

Our implementation uses open or low-cost hardware and we are releasing as open-source software the entire pipeline, thus facilitating reproduction and dissemination.

6.3.1 *General Approach*

We conceived a toolkit to assist the creation of representations of inner activities – our body at large and the hidden processes of our minds in particular, making it visible to oneself and to others. The different components are highlighted in Figure 52. The first step consists in choosing a metric, e.g. the arousal level. For each given metric, there are different ways to measure it. This includes a combination of one or multiple sensor(s) and signal processing algorithm(s). One chooses a support to express this metric (e.g. tangible avatar, screen, speaker for sound) and creates a shape associated to it (e.g. a circle with a changing color, a rhythmic tone). The conjunction of both the shape and the support produces the feedback. It is an iterative process because when one acknowledges the feedback, it changes one's self-representation. Moreover, it creates a feedback loop which affects one's biosignals.

In order to help users mold the system to their likening, we identified three different degrees of freedom:

- The measured physiological signal or mental state (*Metric*)
- The form factor (*Support*)
- The display of the signals (*Shape*)

6.3.1.1 *Sensors and Signal Processing*

Sensors are the hardware used to capture the raw signals of the body. These encompass heart measures such as electrocardiography (ECG) and photoplethysmography (PPG), brain activity measured by electroencephalography (EEG), electrodermal activity (EDA, i.e. perspiration), etc. Once the raw data is acquired, it needs to be processed in order to produce any relevant metric. As an example, heart rate variability can be inferred from the combination of a set of electrodes

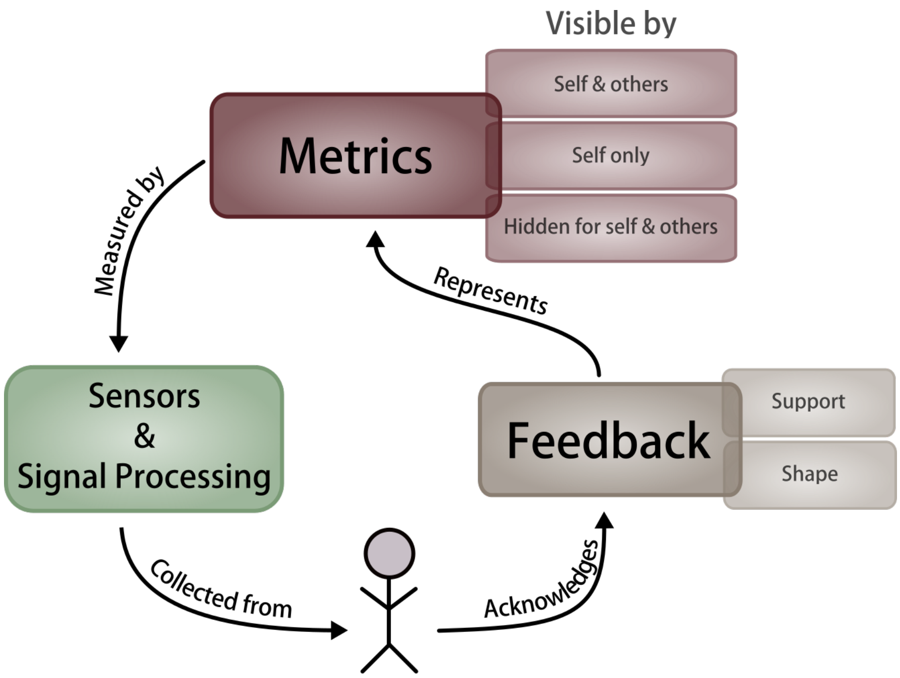


Figure 52: Simplified overview of the toolkit that supports Tobe.

attached to the chest and a QRS wave detector. Emotions can be inferred from EDA or certain frequency bands within EEG.

6.3.2 *Metrics*

There is a continuum in the visibility of the signals and mental states measured from physiological sensors, i.e. metrics. We categorized those metrics in three different levels, depending on who can perceive them without technological help.

1. Perceived by self and others, e.g., eye blinks. Even if those signals may sometimes appear redundant as one may directly look at the person in order to see them, they are crucial in associating a feedback to a user.

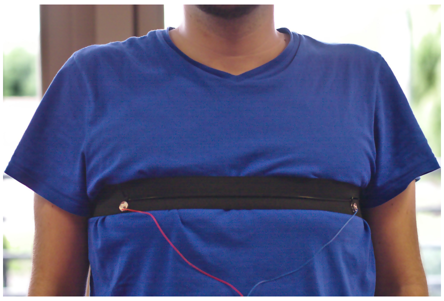
2. Perceived only by self, e.g., heart rate or breathing. Mirroring these signals provides presence towards the feedback.

3. Hidden to both self and others, e.g., mental states such as cognitive workload. This type of metrics holds the most promising applications since they are mostly unexplored.

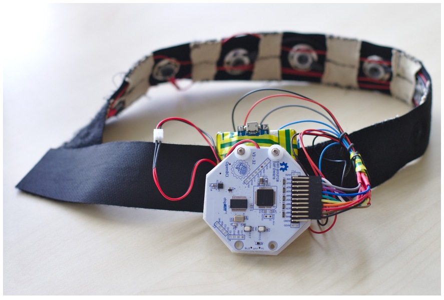
Lower levels (1 & 2) help to breath life into a proxy used to mediate the inner state of the user. These metrics are accessible to our conscious selves. On the other hand, level 3 metrics are little known and are hard to conceptualize for the general public [113] and would benefit the most of a system enabling their visualization.

(a) Coat embedding ECG sensors

(b) Fingerless glove measuring EDA



(c) Breathing belt



(d) EEG headband

Figure 53: Wearable sensors

6.3.3 *Sensors and Signal Processing*

Metrics were acquired from five physiological signals. We measured thoracic circumference for breathing, ECG for heart rate, EDA for arousal, electrooculography (EOG, eyes activity) for eye blinks, and electroencephalography (EEG, brain activity) for most high-level mental states (vigilance, workload, meditation, valence).

We created the sensors with a wearable form factor in mind. Since we used Tobe in public settings, it was important that the sensors were non-invasive (no need to remove clothes or apply gel to the skin) and be quick to install and remove, while being able to acquire a reliable signal. With the setup described in this section, we were able to equip the users and record physiological signals in less than two minutes.

The different sensors were embedded inside a lab coat (Figure 55) which could be put on quickly over daily clothes. This form factor provides enough room in the sleeves and the pockets to take care of the wiring and electronic components storage. The recording of the low-level physiological signals (i.e. everything except EEG) is done using the BITalino board, an Arduino-based recording device. It contains modules that amplifies various physiological signals and embeds a Bluetooth adapter as well as a battery to work in ambulatory settings. Each physiological signal or mental state index was sent to the other stage of the toolkit using LSL[5], a network protocol dedicated to physiological recordings.

6.3.3.1 *ECG*

We chose to use ECG for heart rate activity as it is more accurate than light emission-based methods to detect individual heartbeats [80]. Existing solutions for ECG require electrodes to be put directly on the chest, e.g., heart rate monitor belts. We instead opted for installing TDE-201 Ag-AgCl electrodes from Florida Research Instruments on both wrists of the user (ECG needs two electrodes diametrically opposed to sense heart electrical activity). The electrodes were attached to an elastic band sewed inside the end of the lab coat sleeves which could be tightened with velcro straps (Figure 53a). ECG was recorded with the dedicated ECG module of the BITalino.

6.3.3.2 *EDA*

We measured arousal – which relates to the intensity of an emotion and varies from calm to excited (e.g. *satisfied* vs *happy*) – using EDA. When measuring EDA, most accurate readings can be obtained from the tip of the fingers. However, since it is difficult to manipulate a tangible interface while having hardware attached to one's fingers, we acquire the signal from the palm of a single hand instead. We assess skin conductance from two small conductive thread patches sewn inside a fingerless glove (Figure 53b). Because the BITalino EDA amplifier was not sensitive enough for signals acquired from the palm we made our own, replicating the schematics described in [125].

6.3.3.3 *Breathing*

For breathing, we built a belt based on a stretch sensor (Figure 53c). A conductive rubber band was mounted as a voltage divider and connected to an instrumentation amplifier (Texas Instruments INA128). As opposed to piezoelectric components, that are sensitive to momentous speed instead of position, stretch sensors can directly map users' chest inflation onto their avatar.

6.3.3.4 *EEG and Eye Blinks (EOG)*

We built our own EEG helmet based on the open hardware OpenBCI board (Figure 53d). To shorten setup time we used dry electrodes – the same TDE-201 as for ECG for the forehead, and elsewhere TDE-200 electrodes, which possess small protuberance that could go through the hair. Using a stretchable headband, we restrained electrodes' locations to the rim of the scalp to avoid difficulties with long-haired people. In the 10-20 system, electrodes were positioned at O1,

5 https://github.com/sccn/labstreaminglayer/

P7, F7, FP1, F8, T8, P8 and O2 locations – reference at T7, ground at FP2. We used OpenViBE[6] to analyze physiological data in real time. EEG signals were re-referenced using a common average reference. Mentioned frequencies were extracted with a band-pass filter, taking the log of the power of signals in order to normalize indices. Eye blinks were detected when the signal, after DC drift removal, exceeded 4 times the variance in the F8 electrode. We used the following metrics:

VIGILANCE: appoints for the ability to maintain attention over time. We use the ratio between beta frequency band (15-20Hz) and theta + low alpha frequency band (4-10Hz) for all electrodes [117].

WORKLOAD: increases with the amount of mental effort required to complete a task. We use the ratio between delta + theta band (1-8Hz) in near frontal cortex (F7, FP1, F8, T8) and wide alpha band (8-14Hz) in parietal + occipital cortex (P8, P7, O2, O1) [2, 140].

MEDITATION: we used instantaneous phase locking value between front (FP1, F7, F8) and rear (O1, P7, P8) parts of the brain in alpha + beta bands (7-28Hz) [38] – mindfulness and body focus practices decrease the synchronization while transcendental practice increases it.

VALENCE: designates the hedonic tone of an emotion and varies from positive to negative (e.g., *frustrated* vs *pleasant*). We use the ratio between the EEG signal power in the left (F7, P7, O1) and right (F8, P8, O2) cortex in the alpha band (8-12Hz) [105].

In earlier iterations of the system we tested the use of an Emotive Epoc headset to account for brain activity. The Epoc is a consumers-oriented EEG device, easier to install than medical headsets that use gel. However, it still requires a saline solution that tends to dry over time, causing additional installation time between users. Moreover, good signal quality was next to impossible to obtain with long haired persons. Another downside of consumers-oriented EEG headsets is that they usually conceal signal processing behind proprietary algorithms, with little scientific evidence on what is truly measured. While building a tailored EEG helmet, we took the upper hand on the whole pipeline. With access to raw EEG signals, we looked into the literature to match the inner state we wanted to measure with actual neurological markers.

6.3.4 *Feedback*

The feedback consists in both a support and a shape.

6 http://openvibe.inria.fr/

Figure 54: Modular 3D printed body pieces in order to quickly create a personalized Tobe.

6.3.4.1 *Support*

3D printing a tangible avatar is a powerful incentive for customization. While the version of the system that we deployed in the scientific museum used an already modeled and 3D printed incarnation of Tobe because of time constraints, a user of the system could change the parametric model in order to obtain an avatar that pleases her. The process would be similar to how the appearance of a Nintendo "Mii" can be tuned, except for the tangibility. As a tradeoff between preparation time and customization, we prototyped a "Mr. Potato Head" version of Tobe, with parts ready to be assembled (Figure 54).

Alternatively, it could be possible to resolve some of the parameters of the parametric 3D model based on physiological readings or body properties. Although potentially uncomfortable in social settings, the Body Mass Index of the user could be retrieved from the combination of a video camera and a balance and propose body proportions similar to the person's silhouette. If the user wishes the create an avatar based on inner states, it could for example be possible to alter the stance of the avatar based on the user's real-time stress level. This would allow a dynamic and playful exploration of inner states even at the avatar creation stage.

6.3.4.2 *Shape*

The visualization of users' signals are displayed onto Tobe using Spatial Augmented Reality (SAR), as described in Chapter 2. Despite external hardware requirements – i.e. a projector and eventually a tracking device (Figure 55) – SAR is an easy solution to prototype a system, faster to deploy than putting actual screens in users' surroundings. The augmentation occurred within vvvv[7], a software that

7 http://vvvv.org/

Figure 55: In a scientific museum, various activities were proposed to visitors in order to prompt self-investigation. The setup consists of a projector handling the augmentation and an OptiTrack for the tracking.

uses real-time visual programming to render 3D scenes. As for hardware, we used an LG PF80G projector of resolution 1920x1080 and the tracking of Tobe was achieved with an OptiTrack V120:Trio, running a 120 FPS with an overall latency of 8.3ms and a precision of 0.8mm. The projector was calibrated with the OptiTrack using OpenCV's camera calibration function.

As an alternative to SAR, Tobe can be embedded with small screens, LEDs, actuators and small electronics components so that it represents a standalone unit. We already have a proof of concept of such an implementation thanks to the easiness and accessibility of the building blocks that go with the Arduino platform and the Raspberry Pi (Figure 56).

6.3.4.3 *Customization*

We conceived a GUI that let users draw a picture and animate it according to their wishes. The animator is touch based; users press a "record" button and animate the picture with gestures (Figure 57). Once done, the animation's timeline is automatically mapped to the chosen signal. This animator is kept simple on purpose, it is designed for novice users and as such must remain easy to understand and operate for someone not familiar with animation. Only three basic operations are currently supported – scaling, rotation and translation – and yet it is sufficient to generate meaningful animations. For example, scaling makes a heart beat, translation moves a cloud along respiration and rotation spins cogs faster as workload increases. An advanced tool such as Photoshop has already been integrated as a proof of concept, but the simplicity of the current GUI does not im-

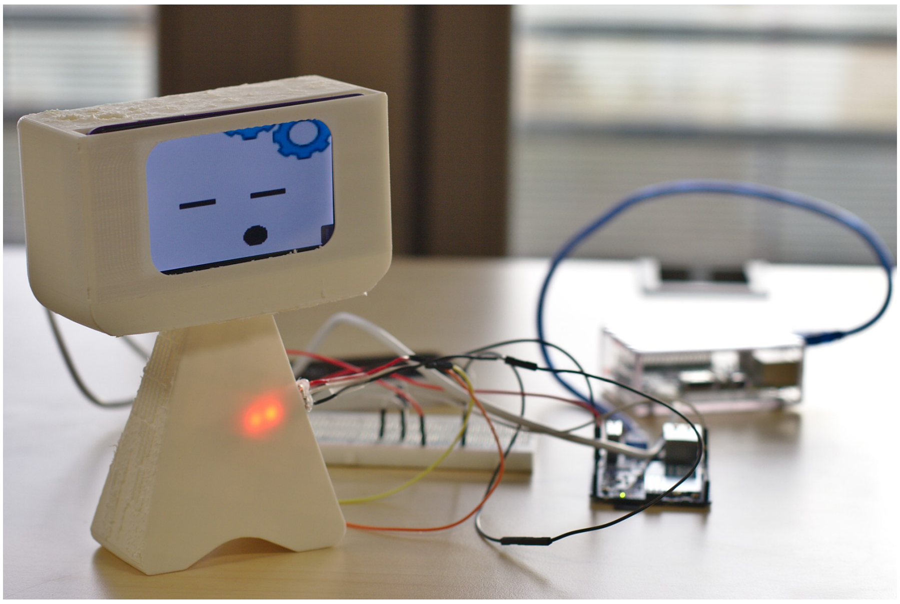


Figure 56: A Tobe with embedded electronics enables a standalone unit that does not require a projector.

pede users' creativity and already is sufficient to enable a tailored feedback.

6.4 TOBE IN THE WILD

We used and tested Tobe in two different applications cases: as a demonstration in a scientific museum and as a multi-user biofeedback device for relaxation and empathy.

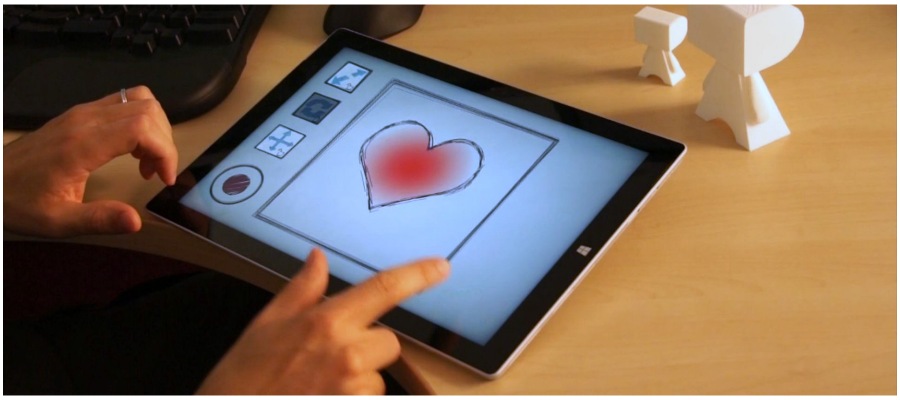


Figure 57: Multitouch animator that let users draw a picture and animate it. The live biosignal of their choice is then used as a way to playback the animation.

6.4.1 *Tobe in a Public Exhibition*

Following a co-design approach, we intervened in a scientific museum over two half days, proposing to passersby to try out Tobe. We built the sensors and prepared the signal processing beforehand because these steps require hardware and expertise. Five high-level metrics were selected: workload, vigilance, meditation, valence and arousal. These metrics were chosen because the general public showed interest into them (meditation and emotions) or because they could benefit from being better known (workload and vigilance). Due to the short duration of our exhibitions, we also set the corresponding feedback (both support and shape), according to the outcome of the questionnaires about people's representations.

After we equipped participants, we gave them "activity cards", a collection of scenarios that were likely to modify their inner state and that prompted self-investigation (Figure 55). There were riddles, arithmetic problems, cute and *less* cute images, a breathing exercise and a "Where's Waldo?" picture. Implicitly the activity cards targeted in this order workload, valence, arousal, meditation and vigilance, but participants were free to test whatever they wanted. These sole cards sufficed to engage participants for a few tens minutes without our intervention.

We created the activity cards after our first intervention in the museum. There were some candies left at disposal next to Tobe to lure museum's visitors to our booth. At some point, one user wanted to see how different tastes affected the emotional valence that was displayed on Tobe. This proved to be a fun activity for him – and for the people around. Having such goal in mind was an effective way to drive participants.

Control of the system was given to users by the mean of a graphical interface (see Figure 58). They had to manipulate the GUI on a nearby tablet to drag and drop visualizations on predefined anchor points. Users could customize some of Tobe's aspects (eyes and heart rate feedback) and among the five high-level metrics available, they selected which one to study at a particular time. When at first we tested Tobe with no control mechanism – i.e., all metrics were displayed altogether – we realized that users were too passive and quickly overwhelmed. The GUI helped to focus and engage users.

To further engage users, Tobe was tracked and participants were asked to put Tobe on a spotlight to "awake" it – i.e., to start physiological signals' streams. The action of bringing life to an inanimate puppet goes well with making the world "magical" again [134], that is to say to use the power of abstraction of modern computer science in order to bring back awe. The aim is not to take benefit of ignorance but to strengthen the amazement that technology can offer. We were ourselves pleasantly disturbed and surprised when we happened to

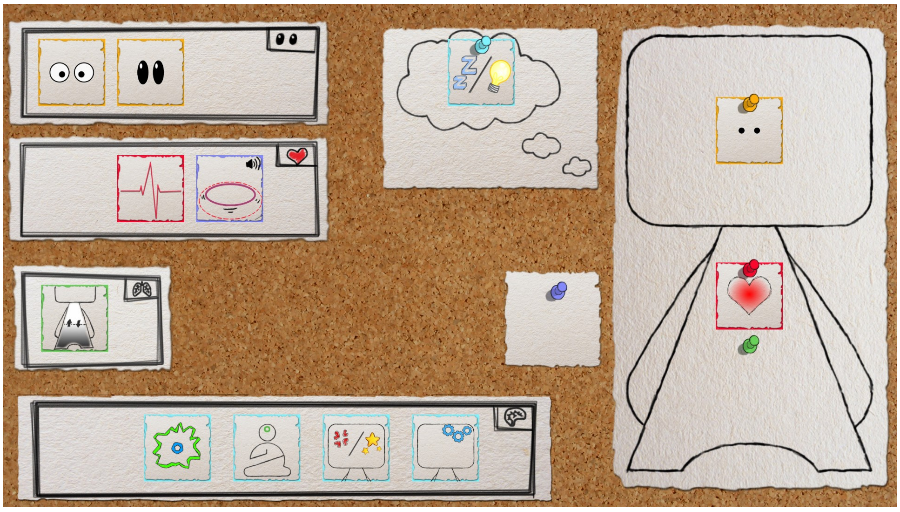


Figure 58: The touch interface allowing visitors to customize the appearance of Tobe's eyes and to select which feedback was displayed on Tobe.

hold in our hands a representation of our beating heart during some routine test. Suddenly the relationship with the digital content felt different, truly tangible.

6.4.2 *Tobe for Multi-Users Relaxation*

We tested Tobe as a relaxation device for two users (Figure 59). The objective was to see if Tobe could be used both as a biofeedback tool and for collaboration.

6.4.2.1 *Implementation*

This version of Tobe relies only on respiration and heart rate variability. It relates to cardiac coherence: when someone takes deep breaths, slowly ($\approx$ 10s periods) and regularly, her or his heart rate (HR) varies accordingly and the resulting state has a positive impact on well-being [101]. During cardiac coherence, HR increases slightly when one inhales and decreases as much when one exhales. We took the magnitude squared coherence between HR and breathing signals over 10s time windows as a "relaxation" index.

Sensors consisted in a breathing belt and in a pair of elastic bands around the wrists to measure ECG. We used OpenBCI instead of BITalino to measure ECG and breathing in order to get more accurate readings. Indeed, the OpenBCI amplifier has a resolution of 24 bits instead of 10 for the BITalino.

There were two Tobes on the table, one for each participant. They were standing at fixed positions, hence no tracking was required. Breathing activity was pictured with inflating lungs onto the torso; cardiac coherence with a blooming flower onto the forehead. The

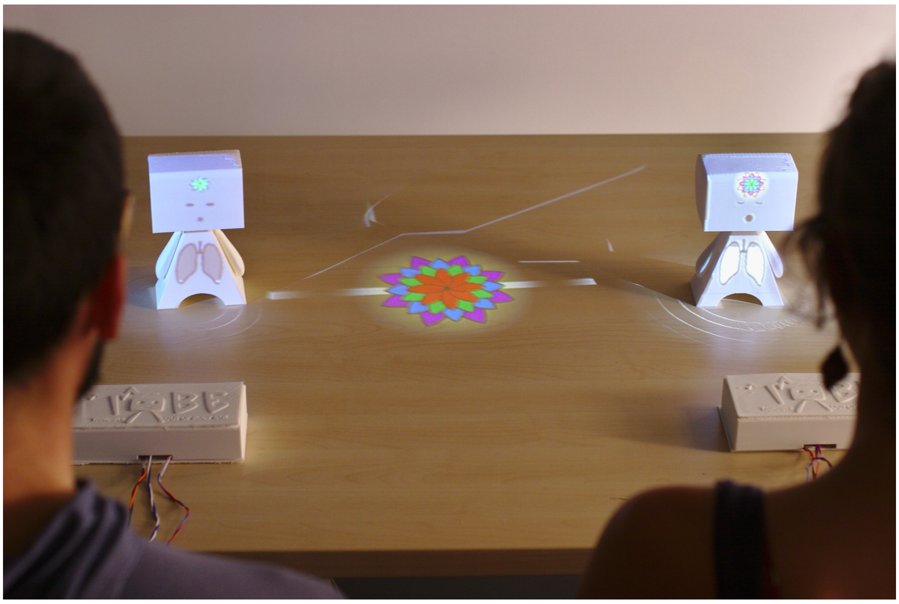


Figure 59: Two users synchronizing their heart rates through cardiac coherence, using Tobe as a biofeedback device.

synchronicity *between* subjects – users' heart rates varying at the same pace – was represented with a similar but bigger flower projected between both Tobes. Additionally, "ripples" on the table, around Tobes' feet, matched heart beats.

6.4.2.2 *Protocol*

We asked 14 participants, by pairs, to come and use Tobe to reach cardiac coherence – 6 females, 8 males, mean age 25.3 (SD=2.8). Participants were coworkers from the same research institute and already knew each other. Participants were seated on each side of a screen and instructed to not talk to each other. We presented them the cardiac coherence activity as a relaxation exercise. Afterward, we equipped them and turned the system on.

The experiment comprised of three sessions of 5 minutes. During the first session, participants had to individually learn how to reach cardiac coherence. A smaller screen on the table prevented them to see each other's Tobe. They had to imitate the breathing pattern given by a gauge going up and down in 5s cycles onto Tobe's body. The lights of the room were dimmed to facilitate a relaxation state and each participant was given headphones playing back rain sounds.

After the training session, the screen separating the two Tobes was removed. Participants were then instructed to repeat the same exercise as before, but without the help of the gauge. They could see their colleague's Tobe. However, there was no interaction between them at this stage – it served as a transition between a self-centered task and a collaboration task.

During the third session, participants were instructed to synchronize their *hearts*. In order to do so, they had to both reach cardiac coherence while breathing on the same rhythm – with no other way to communicate than using their Tobes.

After this final session, we gave questionnaires to participants and conducted informal interviews with them to gather feedback about their experience.

6.4.2.3 *Results and Discussion*

From the questionnaires, that took the form of 5-points Likert scales, participants reported that they were more relaxed after the end of the session: 4.36 on a scale ranging from 1 "much *less* relaxed" to 5 "much *more* relaxed" (SD=0.74). Beside the fact that Tobe acted as an effective biofeedback device, the experiment was also a chance to introduce participants to activities centered around well-being, as few of them were practicing relaxation or meditation in their daily life – 1.93 score (SD=1.44) with 1 "never" and 5 "regularly".

During the interviews, the participants reported that they appreciated the feedback, saying that it formed a coherent experience – e.g. ripples on the table and sounds of rain. Among the few that were practicing yoga regularly, one praised how Tobe favors learning-by-doing over wordy and disrupting instructions but had troubles to follow the 10s breathing cycle since it differed from his usual practice. We had mixed reviews about the visualization associated to breathing, mostly due to the mapping between Tobe's lungs and the measured thoracic circumference being dynamically adapted over time rather than calibrated per user with a min/max. Because of that, some users had to draw their attention away from the breathing patterns in order to achieve cardiac coherence. This is coherent with the results of Wongsuphasawat et al. [189] who observed that visual cues led to more respiratory change and less subjective calm than auditory cues. In hindsight, the respiratory guide should have been mapped to a sound while we could have kept the visual feedback for the actual users' breathing. These two last issues could be resolved by giving users access to the signal processing through our toolkit.

We received comments about how a qualitative and ambient feedback (blooming flower) fostered a better focus on the activity compared to the use of quantitative metrics which are an incentive for competition. Indeed, apart from some comparisons made during the second session, participants did use their Tobes for collaboration. Users described how they use the respiration of their partner to get in sync during the third stage – usually by waiting before inhaling. One participant described how she tried to "help" her companion when he struggled to follow. Another retold how she quickly resumed her regular breathing when she saw that a brief hold troubled her colleague. More playful, a participant laughed afterward at how he purposely "tricked" twice his partner. Even with a communication channel as basic as the display of thoracic circumference, rich interactions emerged between participants over a short period – 5 minutes that felt like less for many of them.

Overall these findings suggest that Tobe could be employed as a proxy for interpersonal communications and that it may have an interesting potential for enhancing well-being.

6.5 APPLICATIONS

We drew usages for Tobe by exploring different dimensions: on the one hand the number of users, Tobes and external observers involved, and, on the other hand, the time and space separating the feedback and the recordings.

6.5.1 *One User*

Tobe can be used as a biofeedback device with a specific goal – e.g., reduce stress – or to gain knowledge about oneself. A feedback about workload and vigilance would prevent overwork. Insights gathered from an introspection session with Tobe could also be employed to *act* better. For example, it might be useful to realize that you are irritated before answering harshly to beloved ones.

6.5.2 *One User and Observer(s)*

Scientists involved in stroke rehabilitation research suggested that Tobe could be used in a medical context. Indeed, patients with motor disabilities may regain mobility after long and difficult sessions of reeducation. However, occasional drawbacks may create anxiety and a counterproductive attitude towards therapy, which leads to even more anxiety. A Tobe could help patients and therapists acknowledge this affective state and break this vicious circle. Autistic persons could also benefit from using Tobe since it is difficult for them and their relatives to gauge their inner state. Explicit arousal could help their integration into society. An offline experiment – i.e. after signals were recorded – pointed to this direction [52].

6.5.3 *Multiple Users and Tobes*

Using Tobe as an alternate communication channel during casual interactions would help to explore connections with relatives, discover and learn from strangers or improve collaboration and efficiency with coworkers. This has been partially explored through the "Reflect Table", which gives a feedback about the affective state of meeting participants [7]; and a bicycle helmet that displays the heart rate of the wearer to the other cyclists nearby has been proposed to support social interactions during physical efforts [173].

6.5.4 *Archetype of a Group*

Tobe could summarize the state of a group. A real-time feedback from the audience would be a valuable tool for every speaker or performer. To pace a course, a teacher could use one Tobe as an overall index that aggregates the attention level of every student in the classroom. Through behavioral measures and with a feedback given afterward, this was investigated in [126].

6.5.5 *Time*

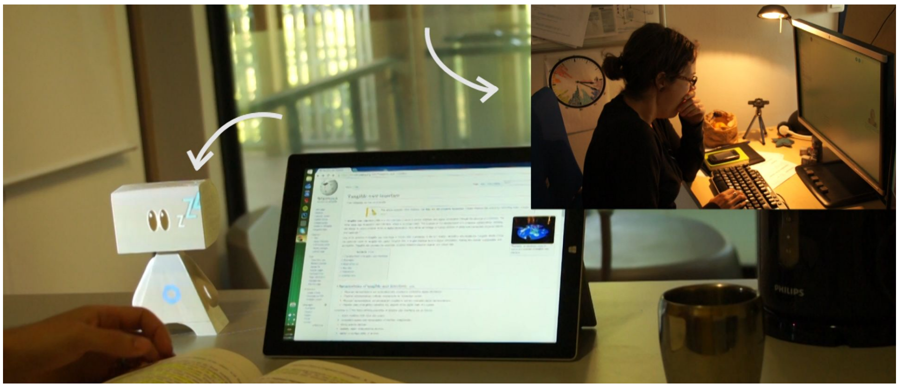
One could want to analyze or to recall inner states after an event. Tobe could be used as a visualization tool for physiological data, the same way as one record video footage of important moments. Instead of replaying what you could see with your eyes, you could visualize how you actually *felt* at that moment in time. Moreover, it could be possible to display both data sources – physiological and video – at the same time, in a synchronized way (Figure 60a). As an example, playing back a video of your wedding would gain an important and very humane dimension.

Collecting large amounts of data – such as is often the case in the QS movement – is useful to gain insight on patterns happening over long stretches of time. However, charts and diagrams are not well suited to convey sensations and feelings. Both graphs and embodied representations are complementary and insight is often created when doing back and forth in abstraction levels [166].

6.5.6 *Space*

If it is not time but space that separates a Tobe from its owner, imagine a distant relationship where the Tobe on your desk slowly awakens as the sun rises in the timezone of your beloved one – and you would wait for Tobe's vigilance to increase to a sufficient level before you pick up your phone for a chat, knowing that your soul mate is a bit grumpy at the beginning of the day (Figure 60b). Besides this theoretical view, it has been hinted that even low-level physiological signals could enhance telepresence [87].

(a) Tobe as a support for playing back recorded internal states. They can be synchronized with a video to view and reflect on cherished memories.



(b) Tobe as an ambient display of the inner state of loved ones far away.

Figure 60: Potential applications for our toolkit.

6.6 SUMMARY

We have presented an open system aimed at externalizing physiological signals and mental states in order to offer users a shared "out-of-body experience". This system covers the entire pipeline, from signals' acquisition to their visualization. Its open nature may be used to introduce STEM discipline to the general public through inquiry-based learning, while end usages can steer them to cognitive science, psychology and humanities, bridging the gap between "hard" and "soft" sciences. Even if the modules we chose promote the inclusion of novices – e.g., visual programming that could be easily extended in OpenViBE or vvvv –, they can be switched to other components that would better suit more experienced users – e.g., Matlab for signal processing. The system is not reduced to a set of tools, though, and we emphasized how such device is aimed at knowing better ourselves and others.

We put the focus on one implementation of the system that consists in a tangible puppet, Tobe, onto which signals are displayed. Its anthropomorphic shape eases users' identification, improves readability and enhances likability. We tested how Tobe affected positively social interactions in a 2-users scenario centered around a relaxation activity. Our co-design approach relied on two interventions that occurred in a scientific museum, as well as on surveys assessing how people relate to physiological signals and how they represent themselves various mental states.

We have identified design dimensions that we used to propose potential applications for our system. Supporting rehabilitation in medical care and facilitating the life in society of individuals with sensory challenges such as autism or ADHD – with the possibility to include the therapists in the loop – are use cases that could benefit from a friendly way to expose inner states. Moreover, it would be interesting to investigate how such a system could be used to ease social interaction and collaboration or to foster empathy towards others.

Future work will include testing Tobe in classrooms or public workshops where users will be invited to build their own self-representation from the ground up, including the tangible support, sensors and the feedback design. Longer usages of the toolkit, over multiple days or weeks, will also be the opportunity to strengthen signal processing in order to provide more reliable mental states that could be displayed between users. Giving users the tools and manuals to investigate their own bodies and mind is a good way to empower them and prompt self-reflection.

Part IV

CONCLUSION

7 CONCLUSION AND FUTURE WORKS

This thesis explored the use of tangible objects as hosts for digital content, added via spatial augmented reality, in a variety of contexts. More specifically, we have focused on the interaction with the digital content as well as using these augmented objects as a basis to foster introspection. Each chapter thus far presented and discussed their individual contributions. In this chapter, we summarize these contributions and discuss potential future works.

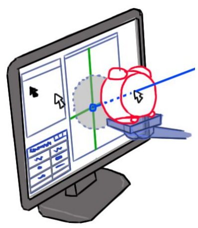
7.1 SUMMARY OF CONTRIBUTIONS

The different work presented in this thesis has focused on two different areas related to tangible augmented objects. First, it explored the *interaction* with the digital content. We have approached this investigation using traditional 2D indirect input devices. While it may at first seem contradictory with an overall tangible approach, our motivation was to build upon already existing – and powerful – digital systems, which digital artists and professionals already spent years learning, and enrich them with physical objects. In a second time, we used these augmented objects as ways to encourage *introspection* activities involving user's internal body states. Overall, we have investigated techniques and created systems that aimed to go beyond traditional screens, involving the physical space around our working environments as well as our own bodies.

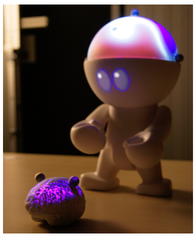
This thesis presented the following four contributions:

1. A study on the performance of a pointing task using 2D indirect pointing devices in a SAR condition compared to a screen condition (Chapter 3). It replicated the interaction technique of a standard desktop computer to a real world scene to see if interacting with real world object in that way could be an appropriate choice. We found that, under certain conditions, pointing in SAR without a screen is still possible (completion time decreased by 11% in SAR in comparison to using a screen). Pointing in SAR still remained predictable using Fitts' law.

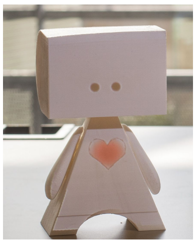
2. A system enabling the seamless interaction between standard screen-based applications and physical objects (Chapter 4). This took the form of an on-screen window – that we called Tangible Viewport – which enables the mouse cursor to slide back and forth between the screen and a physical object held before it. We also explored the interaction space provided by this interplay



between traditional screens – and the applications hosted on them – and tangible augmented objects.



3. A Tangible ElectroEncephaloGraphy Interface – Teegi – (Chapter 5). This is the first system enabling novices to easily visualize and interact with EEG signals in an easy and friendly way. We have shown its use for learning about different brain processes, using the user's own real-time EEG readings as a basis. The interface is controlled through mini-teegis – small tangible puppets – which apply specific filters in order to investigate specific brain activity: motor activity, vision and meditation.



4. A toolkit for creating tangible avatars displaying the real-time low-level physiological signals as well as different mental states of the users (Chapter 6). This set of tools – named Tobe (Tangible Out-of-Body Experience) – encompasses a complete workflow: from the fabrication of the sensors up to the creation and animation of the different visualizations, including signal processing and the physical avatar creation. We created an instance of Tobe which took the shape of a tangible character that was animated according to a user's mental states and physiological responses – e.g. frustration, attention, arousal, heart rate, breathing, ...) – which has been tested in a scientific museum and in a collaborative relaxation task.

7.2 FUTURE WORKS

We conclude this thesis by presenting future works to the different projects discussed in the previous chapters.

7.2.1 *Tangible Viewports Performance Evaluation*

With the Tangible Viewports system (Chapter 4), we have enabled users to interact with tangible augmented objects in a similar way than what they are used to when manipulating 3D objects rendered on screen, with the added benefit of direct manipulation.

The exploratory study we conducted indicated that the metaphor seemed to work well. All participants understood and interacted with the system without showing any sign of being disturbed or bothered by the way of interacting with the augmented objects; every user went on to work on their design without hesitation. However, it would be interesting to quantify the performance of the pointing technique compared to pointing on a viewport displaying a 3D object on screen to assess the impact of physicality. Moreover, it would also be worth comparing the performance of Tangible Viewports with the study conducted with CurSAR (Chapter 3). If a lower decrease in completion time performance is found, it might mean that the implicit reference frame created by screen bezel is important. From there, we could investigate ways to infer this reference frame when using indirect pointing techniques in the real world as it was done in CurSAR.

7.2.2 *Away from the Desktop, Towards the Desk*

With CurSAR (Chapter 3) and Tangible Viewports (Chapter 4), we presented a working environment that leveraged the traditional ways to interact with digital information and combined it with real-world objects and environments. More specifically, Tangible Viewports is an attempt at reducing the gap between the world of screen-based softwares and tangible objects. However, this approach has some limitations. The most obvious one is the constraint of having to position the physical object in front of the screen in order to interact with it using the mouse cursor. Following the general approach of ubiquitous computing and tangible interaction, we think that fusing the desktop computer inside the desk it is laid on would be a promising direction. Of course, much work has already been done under the form of workbenches such as the pioneering work of Krüger et al. on the Responsive Workbench[81] or Ullmer and Ishii's metaDESK [157]. However, it would be interesting to consider a workbench not as a non-immersive virtual workspace, but as a real world work area onto which tools have both digital and physical purposes. In this section, we present a few ideas that we think could be interesting to pursue in this direction.

7.2.2.1 *Complementary Pointing techniques*

When studying the performance of a pointing task in the real world using indirect pointing device in Chapter 3, we ported the technique that is typically used when interacting with 3D elements rendered on screen. However, other indirect techniques[1] might also be worth considering.

Among the different ways to interact with augmented objects, we would like to investigate the use of an object-bound cursor, e.g. Navidget [47]. As is the case when pointing in 3D from 2D input devices, the 3D space has to be collapsed by one dimension. In Chapter 3, we mapped the 2D motion to a plane placed perpendicularly between the user and the scene and raycasted the 3D scene from there. With an object bound cursor, the raycasting could be made from a bounding surface around the physical object – e.g. a sphere – towards the object. A sketch illustrating this idea is shown in Figure 61. This could prove to be interesting, especially in cases where the augmented object is relatively small and held in the hand of the user.

---

1 The consideration for indirect techniques remains justified for precision and prolonged work purposes.

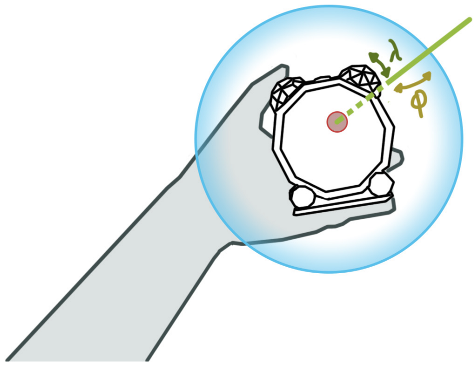
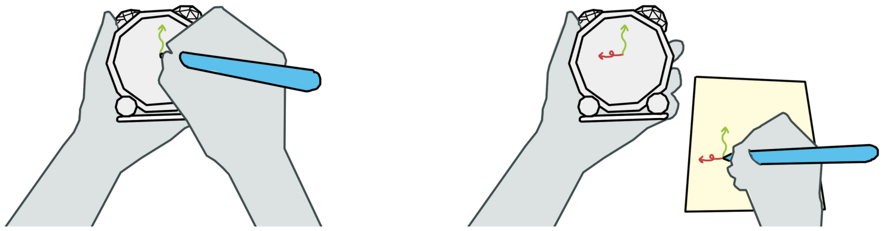


Figure 61: A sketch of a cursor that would be mapped to a bounding volume encompassing the physical object. In this example, the volume is a sphere and the (x, y) mouse coordinates could be mapped to the (φ, λ) coordinate of the sphere, raycasting a tangential ray towards the object.

Another interesting approach would be to combine both direct and indirect methods in a single device. It could take the form of Kienzle and Hinckley's LightRing [77] combined with a Leap Motion or a tracked pen, for example. That way, it would be possible for users to use direct pointing for tasks requiring short operation time and low accuracy while they could simply lay down their dominant hand on the table to move the cursor indirectly from a comfortable position or for a precise operation. Figure 62 depicts the intended principle.

The previous two suggestions are only examples of possible investigation directions for pointing on augmented content hosted on tangible objects. It is important to note that the overall pointing approach when using a SAR system can be very different than one using see-through or video see-through AR. For one, video see-through systems are often being used on multitouch devices, thus offering a surface on which to point. This is effectively making the camera view acts like a 3D rendering window. This enables a variety of pointing techniques that can take place on the multitouch surface [168]. On the other hand, optical see-through devices, such as the Microsoft Hololens[2], rely on gestures and eye tracking. This is because these AR techniques are mostly used to display 3D illusions – things that are "not there". SAR is more limited in this regard: it works best for texture mapping. 3D illusions are still possible, but they require relying on anamorphosis, which tends to not work properly at close distances. However, this means that the digital content is generally hosted on actual physical objects; the location of the virtual objects are close to

2 http://www.microsoft.com/microsoft-hololens/en-us



(a) Direct drawing (b) Indirect drawing

Figure 62: A sketch illustration a combination of direct and indirect interaction for drawing on an object. *(a)* It would be possible to start drawing directly onto the objects surface. *(b)* For doing a more complicated drawing or simply to be able to work on a more stable surface, a user could seamlessly use an indirect cursor handling, using the same input device.

their physical hosted location. This provides opportunities for interaction to be much closer to real world metaphors, such as the one used in TUI techniques.

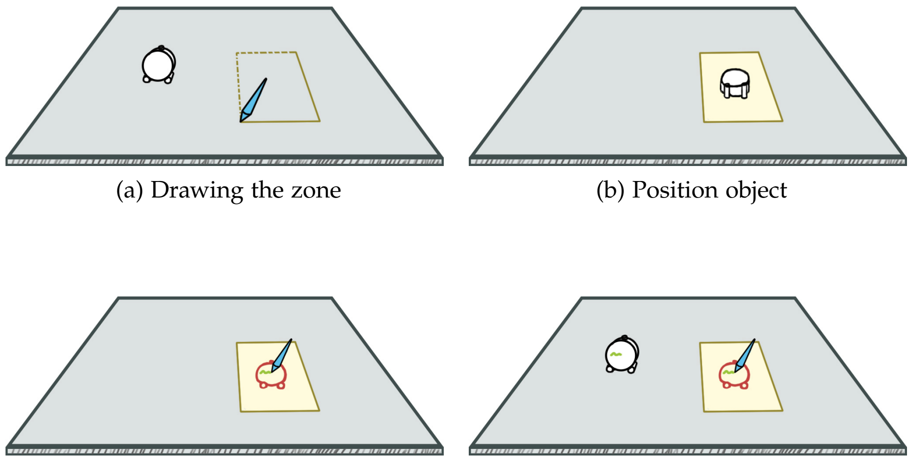
7.2.2.2 *Improvised Interaction Surfaces*

In Chapter 4, we discussed possible screen positions in relation to the augmented object (Section 4.2.1). The choice of handling the object *in front* of the screen was made because we wanted to leverage the use of the native applications of computers which, right now, are operated on screens and also because it helped emphasizing the use of the physical object. However, in an installation where the computer is fused in the desk itself, physical screens might not be permanent fixtures. We could, therefore, imagine a way to improvise working surfaces in order to realize certain operations.

Taking inspiration from the Paper Windows [60] and Tangible Viewports systems, we imagine a workspace where a user could create an interactive window anywhere on the desk using a pen or using a sheet of paper. Then positioning a physical object in front of the newly drawn window and adjusting his head position, he could take a “snapshot” of his view of the physical object. This snapshot would be displayed on the window. He could then easily draw over the object comfortably on a sheet of paper, as it was done in the PapARt system [86]. The appearance of the tangible object would then be synchronized in real time. This workflow is illustrated in Figure 63.

7.2.2.3 *Peripheral Anamorphic Illusions*

We discussed the use of anamorphic illusions in the introductory chapters (1 and 2) and presented the working principle in Section



(a) Drawing the zone
(b) Position object
(c) Work on projection
(d) Content is synchronized

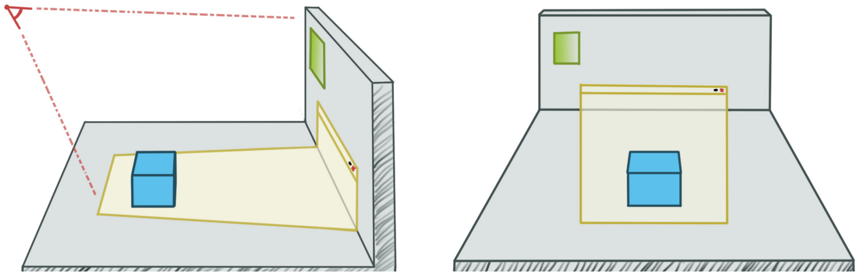
Figure 63: Defining a 2D working surface on the desk and using a tangible object to generate a projective view on which drawing is made easier.

2.2.2. While we did not extensively relied on 3D anamorphic illusions throughout this thesis, we used the principles of perspective in the design of the hybrid widgets (Section 4.3.3.2) within the Tangible Viewports system. While we experimented with the design of these widgets, we quickly realized that having to switch the user's focus from the task to the "illusion" back and forth was being very uncomfortable. Instead, we think that this type of illusions could be leveraged when their purposes are meant to be working in the *periphery* of the user's field of view. It could potentially be used as ambient indicators or guides.

Anamorphic illusions typically do not work well at close distances. This is because of the strong stereo cues – each eye is seeing a slightly different image, enabling the viewer to infer the depth of the objects in sight – when the observed objects are relatively close to the viewer. However, if the user is not focusing on the illusion itself, it might still be possible to perceive it without being uncomfortable[3]. Nevertheless, interesting uses of this principles could prove to be useful in certain interaction contexts.

One example of such use in the context of the work presented in Part ii of this thesis would be in trying to improve indirect pointing in the absence of a physical screen. Since a 2D window – as it was used in Tangible Viewports – acts like a good reference frame for interacting using a 2D indirect input devices, we could project a virtual

3 Of course, if the illusion is in the user's peripheral vision, it will appear blurry. However, it might not be a problem if the intended purpose is to provide peripheral guides to give context and provide subtle cues.



(a) Side-view (b) User's view

Figure 64: A sketch illustrating an anamorphic illusion of a 2D window that can be used as context for indirect pointing techniques in the absence of a screen.

window as a 3D illusion. Since it only acts like context for the cursor behavior, users might not have to put their focus on it in order to reap the benefits of its presence. That way, users might be able to more easily leverage indirect pointing techniques for precision tasks when working with augmented objects. Interestingly, context could be given only at appropriate times, without being distracting when other tasks are being conveyed. Figure 64 presents a sketch of the intended principles.

7.2.3 *Democratization of Physiological Data Visualization*

With Teegi (Chapter 5), one of our objectives was to make complex and abstract data such as EEG signals accessible, not only in terms of expressed meaning but also in ease of manipulation. By displaying the data on a tangible avatar, people were able to explore what was measured via EEG in a friendly and comprehensible way. We already received feedback regarding the potential applications of such ways of displaying EEG information. Among them, we are very interested in the potential for a friendly visualization of rehabilitation progress realized by victims of strokes. Indeed, after a stroke, patients sometime have to go through long periods of rehabilitation that can be frustrating and demoralizing. It would be interesting to study if a system such as Teegi could help these patients be more involved with the recovery process. Being able to see the activity of their brain in real-time as they are training could prove encouraging. Also, comparing with past training sessions would enable patients to assess their progress.

Another possible application we are interested in would be to use Teegi as a support for BCI learning. Indeed, BCI are often difficult to control and require users to be trained in using them efficiently. Usually, BCI rely on mental imagery of motor activities. This skill can be increased through exercises, but the feedback currently used in BCI training is not very helpful [69]. A system such as Teegi could potentially render the training process friendlier while providing feedback to the users.

With Tobe (Chapter 6), we widened the range of physiological signals that we could represent on a tangible avatar. Moreover, we created a set of tools to enable interested users in designing their own avatar and representations for their real-time physiological data. Our goal with Tobe was to democratize even more the access and understanding of physiological data by engaging users in creating their own representations of their inner selves. We already tested the use of the toolkit in a scientific museum, but we want to give the toolkit to a wider range of users – for example in FabLabs and different maker spaces – and see where it can be improved upon and what kind of applications people are interested in. One future perspective we are particularly excited about is the use of such a tangible avatar in medical settings and health applications. Indeed, we want to study if displaying medical readings on a physical puppet could help patients be more involved in monitoring their own states. Moreover, it could also be useful as an ambient display for medical personnel, potentially providing a complementary "at-a-glance" view of the patient's readings. Since the form factor can be customized, we could think of ways to make health monitoring more engaging and positive for kids – by 3D printing avatars shaped as their favorite super heroes, for example.

7.2.4 *Fostering Well-being and Self-Reflection*

One of the main underlying motivation while developing both Teegi and Tobe (respectively Chapter 5 and 6) was anchored in a desire to know more about our inner selves. Of course, physiological computing does not consist in an open door to the mind – it is not possible to read thoughts. However, it consists in an interesting technological mean to gain insight about how our bodies and minds work. Interestingly, while demoing Tobe in a scientific museum and presenting the technology as a mean to know more about ourselves, one of the visitors stayed a moment to discuss it. She was emphasizing the fact that technology was unnecessary for this purpose and that introspection came from deliberate practices like meditation and contemplation. Here, we agree in that introspection is an activity that requires self-motivation. This hardly can be substituted for some technological bells and whistles. However, while not completely necessary, we believe that it can be a powerful tool to *support* and *encourage* meditative, contemplative and self-reflective practices. Some works have started to point in this direction, for example using conscious breathing patterns [106, 107]. We also used this type of exercise with Tobe

(Section 6.4.2). However, much work remains to be done for creating technologies that focus on human wellbeing.

One aspect we are very interested in is the combination of slow technology [48] and physiological computing. In the same way Weiser and Brown envisioned it [178], we think that this combination of fields could present great opportunities for fostering calm and well-being. Even more so, with technology designed for self-reflection, people could investigate how their own inner states are affected by different stimuli and events.

7.3 FINAL REMARKS

As Brown [22] puts it in his talk, the most important aspect of Weiser and Brown's vision [178] of ubiquitous computing might have been overlooked: a technology that would create *calm*. Rogers [133], for example, has instead been advocating focusing on making the user *engaged* with technology. Engagement is indeed an important factor. It makes users involved with what they are seeing, potentially shaping the digital content for themselves and others. However, I sincerely believe that engaged experiences should be interwoven with calm ones.

Each time I find myself isolated from digital technologies for an extended period of time, for example when doing a self-supported bicycle tour of multiple weeks, it always strikes me how I do not miss the digital world much. Never do I wonder what is happening on the internet or if my emails are piling up. The real world is fully engaging me on its own, without being overwhelming. However, upon coming back home, the technology is suddenly retrieving its exhilarating appeal. Even though I feel very stimulated and engaged with the technology, rarely does it make me feel as calm as when I ride my bicycle or sleep in a tent. I think that we shouldn't have to choose between using technology and being calm. It is worthwhile pursuing a way to combine both: how can we use digital technology to sometimes feel engaged and productive while at other times slowing down and be mindful to what matters to us personally?

In this thesis, I have focused on the use of physical objects in conjunction with digital content. I believe, as David Rose puts it in his book on "Enchanted Objects" [134], that merging both the digital and real worlds, with a special emphasis on reality, has the potential for technology to stay closer to our humane nature.

Part V

APPENDIX

A NOTES AND CLARIFICATIONS

A.1 PRECISION AND ACCURACY

Precision and accuracy are two different concepts that are often use as synonyms in a normal discourse. However, in scientific contexts, their difference is important to highlight. The Oxford English dictionary defines the terms as such:

ACCURACY The closeness of a measurement, calculation, or specification to the correct value.

PRECISION The degree of refinement in a measurement, calculation, or specification, esp. as represented by the number of digits given.

In terms of tracking, a precise system would not jitter and an accurate system would provide a pose that is as close as possible to the true pose of the physical object being tracked. Figure 65 visually highlights the difference between the two concepts.

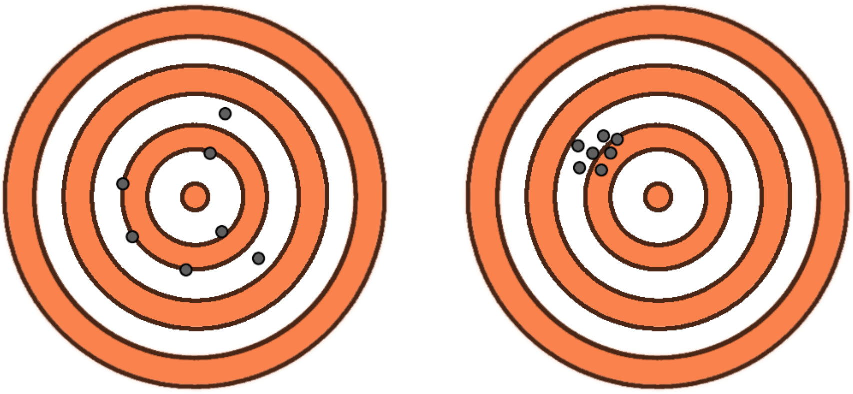
A.2 PINHOLE CAMERA MODEL

The most commonly used camera model to represent the projection of a scene on an image plane is the *pinhole camera model* (also referred as the *perspective camera model*). An interested reader wanting a more in-depth view on this topic is refered to the book of Trucco and Verri [155].

Images are, by definition, mostly in two dimensions. That means that a point in the real world needs to be converted somehow to find its place on the image's projection plane. Between these two coordinate systems, there is the camera world that refers to 3D points from the point-of-view of the camera. This set of transformations includes the *intrinsic* and *extrinsic* camera parameters. The first allows to project a given point expressed in the camera's coordinate system onto the projection plane while the second allows to define the position and orientation of the camera according to the reference coordinate system (often referred as the world's coordinate system).

A.2.1 *Intrinsic parameters*

Trucco and Verri [155] defines the intrinsic parameters as follow:



(a) An accurate but imprecise distribution. (b) A precise but inaccurate distribution.

Figure 65: A visual representation of the difference between *accuracy* and *precision*.

Intrinsic parameters can be defined as the set of parameters needed to characterize the optical, geometric, and digital characteristics of the viewing camera. For a pinhole camera, there are three sets of intrinsic parameters, specifying respectively:

- the perspective projection, for which the only parameter is the focal length, f;
- the transformation between the camera frame coordinates and pixel coordinates;
- the geometric distortion introduced by the optics.

The pinhole camera model is illustrated in Figure 66. The model consists of an image plane π and a 3D point O, representing the *camera center* or the *focus of projection*. The positive z axis that goes from $O = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^t$ through the center $c = \begin{bmatrix} c_x & c_y & f \end{bmatrix}^t$ of the image plane is called the *optical axis*.

The first set of parameters (the focal length f[1]) allows to project a 3D point in the camera world $P = \begin{bmatrix} X & Y & Z \end{bmatrix}^t$ on the image plane π at point $p = \begin{bmatrix} x & y & z \end{bmatrix}^t$[2]. The image plane, perpendicular to the optical axis, is located at a distance of the focal length f of the camera. The following equations allow the projection of P on the image plane π, expressed in the camera coordinate system:

$$x = f\frac{X}{Z}$$

1 The focal length can be different in x and y direction resulting in having two focal length values f_x and f_y. We however make the assumption that these values are equals resulting in a single f value.

2 As the z component of a point on the image plane is always f, the representation $p = \begin{bmatrix} x & y \end{bmatrix}^t$ is used instead of $p = \begin{bmatrix} x & y & f \end{bmatrix}^t$.

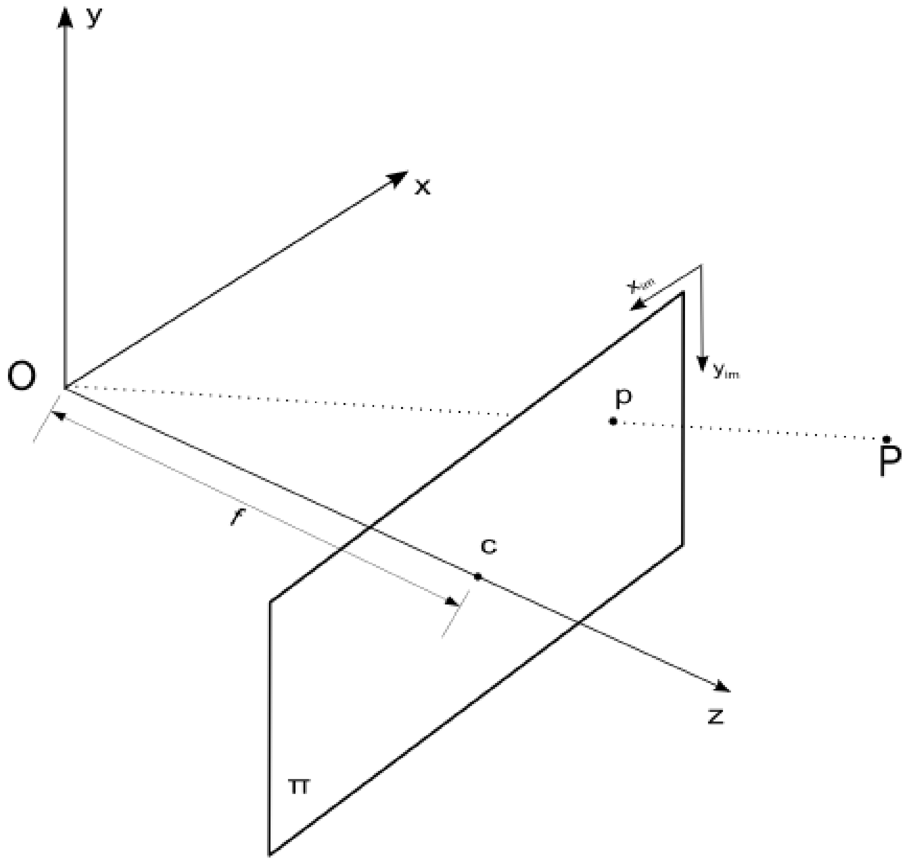


Figure 66: The pinhole or perspective camera model. The model consists of an image plane π and a 3D point O, representing the *camera center* or the *focus of projection*. The positive z axis that goes from O through the center c of the image plane is called the *optical axis*. A 3D point in the world P is projected on the image plane π at point p. x_{im} and y_{im} represents the coordinate system of the image.

$$y = f\frac{Y}{Z} \tag{1}$$

The second set of parameters allows to convert the projected point $p = (x, y)$, expressed in the camera's coordinate system, in pixel coordinates (image's coordinates) (x_{im}, y_{im}). This conversion also involves the actual, physical dimensions (in mm) of the pixels of the camera (s_x, s_y) and the center of the image plane $c = (c_x, c_y)$. Note that the sign change in Equation 2 is due to the fact that the camera's coordinate system and the image reference frame have inverted x and y axis (Figure 66).

$$x = -(x_{im} - c_x)s_x$$

$$y = -(y_{im} - c_y)s_y \tag{2}$$

The third and last set of parameters is related to the radial distortion introduced by the imperfections of the optics in the camera. The distortion becomes really visible near the edge of the image (distortion at (c_x, c_y) is null) – easily resulting in a shift of several pixels. Fortunately, these deformations can be modeled with parameters $k_1, \ldots, k_n$ where, in most standard calibrations, $n = 4$ is sufficient to undistort an image for it to be usable for processing.

Most of CV algorithms expect an undistorted image as an input. It means that the image should be remapped to compensate for its shifted center c and the radial distortion modeled with parameters $k_1, \ldots, k_n$. Most CV libraries, such as OpenCV[3], can determine these sets of parameters using an once-in-a-lifetime calibration process and undistort images.

A.2.2 *Extrinsic parameters*

Extrinsic parameters represents the transformation from the world's coordinate system to the camera reference frame. This can simply be modeled using a rotation R and a translation T in 3D space. A point in the world can be brought in the camera's coordinate system using the following matrix transformation:

$$P_c = R(P_w - T) \tag{3}$$

A.2.3 *Transformations summary*

If a point $P = \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}_w^t$ in the world needs to be converted to a position in pixels in the image's reference frame ($\begin{bmatrix} x & y & z \end{bmatrix}_{im}^t$),

3 http://opencv.org/

all the transformations described above can be combined in a single matrix expression, where M_{int} and M_{ext} are respectively the matrix representing the intrinsic and extrinsic parameters of the camera, f is the focal length of the camera, (s_x, s_y) are the physical dimensions (in mm) of the pixels of the camera, R is a rotation matrix and T is a translation matrix:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{im} = M_{int} M_{ext} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}_{w} \tag{4}$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{im} = \begin{pmatrix} -f/s_x & 0 & c_x \\ 0 & -f/s_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} R_{3x3} \vdots & -R^t T_{3x1} \end{pmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}_{w} \tag{5}$$

BIBLIOGRAPHY

[1] Eric Akaoka, Tim Ginn, and Roel Vertegaal. Displayobjects: prototyping functional physical interfaces on 3d styrofoam, paper or cardboard models. In *TEI*, pages 49–56. ACM, 2010. (Cited on pages 55 and 65.)

[2] Pavlo Antonenko, Fred Paas, Roland Grabner, and Tamara van Gog. Using Electroencephalography to Measure Cognitive Load. *Educational Psychology Review*, 22(4):425–438, 2010. (Cited on page 101.)

[3] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The internet of things: A survey. *Computer Networks*, 54(15):2787 – 2805, 2010. (Cited on pages 4 and 18.)

[4] S. Audet and M. Okutomi. A user-friendly method to geometrically calibrate projector-camera systems. *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 47–54, jun 2009. (Cited on page 32.)

[5] Samuel Audet and Jeremy R Cooperstock. Shadow removal in front projection environments using object tracking. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007. (Cited on page 38.)

[6] Ronald T Azuma et al. A survey of augmented reality. *Presence*, 6(4):355–385, 1997. (Cited on pages 4 and 17.)

[7] Khaled Bachour. *Augmenting Face-to-Face Collaboration with Low-Resolution Semi-Ambient Feedback*. PhD thesis, EPFL, 2010. (Cited on page 110.)

[8] Ravin Balakrishnan and Gordon Kurtenbach. Exploring bimanual camera control and object manipulation in 3d graphics interfaces. In *CHI*, pages 56–62. ACM, 1999. (Cited on page 65.)

[9] D. Bandyopadhyay, R. Raskar, and H. Fuchs. Dynamic shader lamps : painting on movable objects. *ISAR*, pages 207–216, 2001. (Cited on pages 22, 45, and 54.)

[10] Ditte Amund Basballe and Kim Halskov. Projections on museum exhibits: engaging visitors in the museum setting. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*, pages 80–87. ACM, 2010. (Cited on page 38.)

[11] Patrick Baudisch, Edward Cutrell, Ken Hinckley, and Robert Gruen. Mouse ether: accelerating the acquisition of targets across multi-monitor displays. In *CHI'04 extended abstracts on Human factors in computing systems*, pages 1379–1382. ACM, 2004. (Cited on page 45.)

[12] Patrick Baudisch, Henning Pohl, Stefanie Reinicke, Emilia Wittmers, Patrick Lühne, Marius Knaust, Sven Köhler, Patrick Schmidt, and Christian Holz. Imaginary reality gaming: ball games without a ball. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, pages 405–410. ACM, 2013. (Cited on page 18.)

[13] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *Computer vision–ECCV 2006*, pages 404–417. Springer, 2006. (Cited on page 29.)

[14] Hrvoje Benko, Andrew D Wilson, and Ravin Balakrishnan. Sphere: multi-touch interactions on a spherical display. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 77–86. ACM, 2008. (Cited on page 38.)

[15] Hrvoje Benko, Ricardo Jota, and Andrew Wilson. Miragetable: freehand interaction on a projected augmented reality tabletop. In *CHI*, pages 199–208. ACM, 2012. (Cited on pages 39 and 45.)

[16] Hrvoje Benko, Andrew D Wilson, Federico Zannier, and Hrvoje Benko. Dyadic projected spatial augmented reality. In *UIST*, pages 645–655. ACM, 2014. (Cited on pages 19, 21, and 66.)

[17] Emily Bennett and Brett Stevens. The effect that touching a projection augmented model has on object-presence. In *Information Visualisation, 2005. Proceedings. Ninth International Conference on*, pages 790–795. IEEE, 2005. (Cited on page 37.)

[18] Paul J Besl and Neil D McKay. Method for registration of 3-d shapes. In *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, pages 586–606. International Society for Optics and Photonics, 1992. (Cited on page 30.)

[19] Oliver Bimber and Ramesh Raskar. *Spatial augmented reality: merging real and virtual worlds*. CRC Press, 2005. (Cited on page 17.)

[20] Jordi Brandts, Arno Riedl, and Frans Van Winden. Competition and well-being. 2005. (Cited on page 95.)

[21] Eric Brockmeyer, Ivan Poupyrev, and Scott Hudson. Papillon: Designing curved display surfaces with printed optics. In *Pro-*

ceedings of the 26th Annual ACM Symposium on User Interface Software and Technology, UIST '13, pages 457–462, New York, NY, USA, 2013. ACM. (Cited on page 26.)

[22] John NA Brown. Expert talk for time machine session: Designing calm technology" as refreshing as taking a walk in the woods". In *2012 IEEE International Conference on Multimedia and Expo*, page 423. IEEE, 2012. (Cited on page 123.)

[23] William Buxton. Integrating the periphery and context: A new taxonomy of telematics. In *Proceedings of graphics interface*, volume 95, pages 239–246. Citeseer, 1995. (Cited on page 8.)

[24] Jessica R Cauchard, Markus Löchtefeld, Pourang Irani, Johannes Schoening, Antonio Krüger, Mike Fraser, and Sriram Subramanian. Visual separation in mobile multi-display environments. In *UIST*, pages 451–460. ACM, 2011. (Cited on page 59.)

[25] G. Chanel and C. Muhl. Connecting Brains and Bodies: Applying Physiological Computing to Support Social Interaction. *Interacting with Computers*, 2015. (Cited on page 95.)

[26] Marcelo Coelho and Pattie Maes. Shutters: a permeable surface for environmental control and communication. In *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, TEI '09, pages 13–18, New York, NY, USA, 2009. ACM. (Cited on page 10.)

[27] Jonathan Cohen. Defining Identification: A Theoretical Look at the Identification of Audiences With Media Characters. *Mass Communication and Society*, 4(3):245–264, 2001. (Cited on page 91.)

[28] Gillian Crampton Smith. The hand that rocks the cradle. *ID magazine*, pages 60–65, 1995. (Cited on page 9.)

[29] Antonio R. Damasio. *Descartes' error: emotion, reason, and the human brain*. 1994. ISBN 0-399-13894-3. (Cited on page 91.)

[30] Andrew J Davison, Ian D Reid, Nicholas D Molton, and Olivier Stasse. Monoslam: Real-time single camera slam. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(6):1052–1067, 2007. (Cited on page 29.)

[31] Sanne Dekker, Nikki C Lee, Paul Howard-Jones, and Jelle Jolles. Neuromyths in education: Prevalence and predictors of misconceptions among teachers. *Front. Psychol.*, 3, 2012. (Cited on page 78.)

[32] Daniel Dobler, Michael Haller, and Philipp Stampfl. Asr: augmented sound reality. In *ACM SIGGRAPH 2002 conference abstracts and applications*, pages 148–148. ACM, 2002. (Cited on page 18.)

[33] Henry Been-Lirn Duh and Mark Billinghurst. Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 193–202, sep 2008. (Cited on page 27.)

[34] Stephen H Fairclough. Fundamentals of physiological computing. *Interacting with computers*, 21(1):133–145, 2009. (Cited on pages 7, 9, and 94.)

[35] M. Fatourechi, A. Bashashati, R. Ward, and G. Birch. EMG and EOG artifacts in brain computer interface systems: A survey. *Clin Neurophysiol*, 118(3):480–494, 2007. (Cited on page 84.)

[36] George W. Fitzmaurice, Hiroshi Ishii, and William A. S. Buxton. Bricks: Laying the foundations for graspable user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '95, pages 442–449, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co. (Cited on page 9.)

[37] Stéphanie Fleck and Gilles Simon. An Augmented Reality Environment for Astronomy Learning in Elementary Grades: An Exploratory Study. In *IHM '13*, page 9, 2013. (Cited on page 78.)

[38] Jérémy Frey, Renaud Gervais, Stéphanie Fleck, Fabien Lotte, and Martin Hachet. Teegi: tangible eeg interface. In *UIST*, pages 301–308. ACM, 2014. (Cited on pages 54, 95, and 101.)

[39] Jérémy Frey, Christian Mühl, Fabien Lotte, and Martin Hachet. Review of the use of electroencephalography as an evaluation method for human-computer interaction. In *PhyCS '14*, 2014. (Cited on page 94.)

[40] Renaud Gervais, Jérémy Laviole, Asier Marzo, and Martin Hachet. The Good, the Bad and the Hacked: Creative Coding on Objects. Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction Work-in-Progress, feb 2014. URL https://hal.inria.fr/hal-01096299.

[41] Renaud Gervais, Jérémy Frey, and Martin Hachet. Pointing in spatial augmented reality from 2d pointing devices. In *INTERACT*, page 8, 2015.

[42] Renaud Gervais, Jérémy Frey, Alexis Gay, Fabien Lotte, and Martin Hachet. TOBE: Tangible Out-of-Body Experience. In

Proceedings of the 10th International Conference on Tangible, Embedded and Embodied Interaction, TEI '16, Eindhoven, Netherlands, feb 2016. ACM.

[43] Renaud Gervais, Joan Sol Roo, and Martin Hachet. Tangible Viewports: Getting Out of Flatland in Desktop Environments. In *Proceedings of the 10th International Conference on Tangible, Embedded and Embodied Interaction*, TEI '16, Eindhoven, Netherlands, feb 2016. ACM.

[44] James J Gibson. The theory of affordances. *Hilldale, USA*, 1977. (Cited on page 4.)

[45] James Jerome Gibson. *The ecological approach to visual perception.* Houghton, Mifflin and Company, 1979. (Cited on page 4.)

[46] Connie Golsteijn, Elise Van Den Hoven, David Frohlich, and Abigail Sellen. Towards a more cherishable digital object. In *Proceedings of the Designing Interactive Systems Conference*, pages 655–664. ACM, 2012. (Cited on page 95.)

[47] Martin Hachet, Fabrice Decle, Sebastian Knoedel, and Pascal Guitton. Navidget for Easy 3D Camera Positioning from 2D Inputs. In *Proceedings of the IEEE Symposium on 3D User Interfaces (3DUI)*, pages 83–88, United States, 2008. (Cited on page 117.)

[48] Lars Hallnäs and Johan Redström. Slow technology–designing for reflection. *Personal and ubiquitous computing*, 5(3):201–212, 2001. (Cited on pages 8 and 123.)

[49] Richard W Hamming and JF Kaiser. You and your research. 1986. (Cited on page 7.)

[50] Chris Harrison, Hrvoje Benko, and Andrew D. Wilson. Omnitouch: Wearable multitouch interaction everywhere. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, pages 441–450, New York, NY, USA, 2011. ACM. (Cited on page 44.)

[51] Chris Harrison, Robert Xiao, Julia Schwarz, and Scott E Hudson. Touchtools: leveraging familiarity and skill with physical tools to augment touch interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2913–2916. ACM, 2014. (Cited on page 54.)

[52] Elliott Hedman, Lucy Miller, Sarah Schoen, and Darci Nielsen. Measuring Autonomic Arousal During Therapy. In *D&E '12*, pages 11–14, 2012. (Cited on page 109.)

[53] Suzana Herculano-Houzel. Do you know your brain? A survey on public neuroscience literacy at the closing of the decade

of the brain. *The Neuroscientist*, 8(2):98–110, 2002. (Cited on page 78.)

[54] Valentin Heun, James Hobin, and Pattie Maes. Reality editor: Programming smarter objects. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, pages 307–310, New York, NY, USA, 2013. ACM. (Cited on page 64.)

[55] Valentin Heun, Shunichi Kasahara, and Pattie Maes. Smarter objects: using ar technology to program physical objects and their interactions. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 961–966, New York, NY, USA, 2013. ACM. (Cited on page 64.)

[56] Otmar Hilliges, David Kim, Shahram Izadi, Malte Weiss, and Andrew Wilson. Holodesk: direct 3d interactions with a situated see-through display. In *CHI*, pages 2421–2430. ACM, 2012. (Cited on page 55.)

[57] Ken Hinckley, Randy Pausch, John C Goble, and Neal F Kassell. Passive real-world interface props for neurosurgical visualization. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 452–458. ACM, 1994. (Cited on page 79.)

[58] David Holman and Hrvoje Benko. Sketchspace: designing interactive behaviors with passive materials. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 1987–1992, New York, NY, USA, 2011. ACM. (Cited on page 64.)

[59] David Holman and Roel Vertegaal. Organic user interfaces: designing computers in any way, shape, or form. *Commun. ACM*, 51(6):48–55, jun 2008. (Cited on pages 4 and 10.)

[60] David Holman, Roel Vertegaal, Mark Altosaar, Nikolaus Troje, and Derek Johns. Paper windows: interaction techniques for digital paper. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 591–599. ACM, 2005. (Cited on page 119.)

[61] David Holman, Audrey Girouard, Hrvoje Benko, and Roel Vertegaal. The design of organic user interfaces: Shape, sketching and hypercontext. *Interacting with Computers*, 25(2):133–142, 2013. (Cited on pages 4 and 10.)

[62] Michael S Horn, Erin Treacy Solovey, R Jordan Crouser, and Robert JK Jacob. Comparing the use of tangible and graphical

programming languages for informal science education. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 975–984. ACM, 2009. (Cited on pages 79 and 97.)

[63] Eva Hornecker. The role of physicality in tangible and embodied interactions. *interactions*, 18(2):19–23, 2011. (Cited on page 97.)

[64] Hiroshi Ishii. Tangible bits: beyond pixels. In *Proceedings of the 2nd international conference on Tangible and embedded interaction*, pages xv–xxv. ACM, 2008. (Cited on page 9.)

[65] Hiroshi Ishii and Brygg Ullmer. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *CHI*, pages 234–241. ACM, 1997. (Cited on pages 4, 9, and 54.)

[66] Hiroshi Ishii, Craig Wisneski, Scott Brave, Andrew Dahley, Matt Gorbet, Brygg Ullmer, and Paul Yarin. ambientroom: Integrating ambient media with architectural space. In *CHI 98 Cconference Summary on Human Factors in Computing Systems*, CHI '98, pages 173–174, New York, NY, USA, 1998. ACM. (Cited on page 8.)

[67] Hiroshi Ishii, Dávid Lakatos, Leonardo Bonanni, and Jean-Baptiste Labrune. Radical atoms: Beyond tangible bits, toward transformable materials. *interactions*, 19(1):38–51, jan 2012. (Cited on pages 9 and 10.)

[68] Yvonne Jansen, Pierre Dragicevic, and Jean-Daniel Fekete. Evaluating the efficiency of physical visualizations. In *CHI*, pages 2593–2602. ACM, 2013. (Cited on page 66.)

[69] Camille Jeunet, Alison Cellard, Sriram Subramanian, Martin Hachet, Bernard N'Kaoua, and Fabien Lotte. How Well Can We Learn With Standard BCI Training Approaches? A Pilot Study. In *6th International Brain-Computer Interface Conference*, Graz, Austria, sep 2014. (Cited on page 122.)

[70] Steven Johnson. *Where good ideas come from: The natural history of innovation*. Penguin UK, 2010. (Cited on page 7.)

[71] Brett Jones, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, Nikunj Raghuvanshi, and Lior Shapira. Roomalive: Magical experiences enabled by scalable, adaptive projector-camera units. In *UIST*, pages 637–644. ACM, 2014. (Cited on page 71.)

[72] Daniel Kahneman. *Thinking, fast and slow*. Macmillan, 2011. (Cited on page 96.)

[73] Martin Kaltenbrunner and Ross Bencina. reactivision: a computer-vision framework for table-based tangible interaction. In *Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 69–74. ACM, 2007. (Cited on page 28.)

[74] Shaun K Kane, Daniel Avrahami, Jacob O Wobbrock, Beverly Harrison, Adam D Rea, Matthai Philipose, and Anthony LaMarca. Bonfire: a nomadic system for hybrid laptop-tabletop interaction. In *UIST*, pages 129–138. ACM, 2009. (Cited on page 55.)

[75] Hirokazu Kato and Mark Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Augmented Reality, 1999.(IWAR'99) Proceedings. 2nd IEEE and ACM International Workshop on*, pages 85–94. IEEE, 1999. (Cited on pages 28 and 29.)

[76] Rohit Ashok Khot, Larissa Hjorth, and Florian'Floyd' Mueller. Understanding physical activity through 3d printed material artifacts. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 3835–3844. ACM, 2014. (Cited on page 95.)

[77] Wolf Kienzle and Ken Hinckley. Lightring: always-available 2d input on any surface. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, pages 157–160. ACM, 2014. (Cited on page 118.)

[78] John Klein, Michael Stern, Giorgia Franchin, Markus Kayser, Chikara Inamura, Shreya Dave, James C Weaver, Peter Houk, Paolo Colombo, Maria Yang, et al. Additive manufacturing of optically transparent glass. *3D Printing and Additive Manufacturing*, 2, 2015. (Cited on page 26.)

[79] Clemens Nylandsted Klokmose, Janus Bager Kristensen, Rolf Bagge, and Kim Halskov. Bullseye: High-precision fiducial tracking for table-based tangible interaction. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, pages 269–278. ACM, 2014. (Cited on page 28.)

[80] J. Kranjec, S. Beguš, G. Geršak, and J. Drnovšek. Non-contact heart rate and heart rate variability measurements: A review. *Biomedical Signal Processing and Control*, 13:102–112, sep 2014. (Cited on page 100.)

[81] Wolfgang Krüger, Christian-A Bohn, Bernd Fröhlich, Heinrich Schüth, Wolfgang Strauss, and Gerold Wesche. The responsive workbench: A virtual work environment. *Computer*, (7):42–48, 1995. (Cited on page 117.)

[82] Kai Kunze, Mikio Iwamura, Kenji Kise, Seiichi Uchida, and Shinichiro Omachi. Activity recognition for the mind: Toward a cognitive" quantified self". *Computer*, 46(10):105–108, 2013. (Cited on page 94.)

[83] Jean-Philippe Lachaux, Eugenio Rodriguez, Michel Le Van Quyen, Antoine Lutz, Jacques Martinerie, and Francisco J. Varela. Studying single-trials of phase synchronous activity in the brain. *Int. J. Bifurcation Chaos*, 10(10):2429–2439, oct 2000. (Cited on page 84.)

[84] Marc Lambooij, Marten Fortuin, Ingrid Heynderickx, and Wijnand IJsselsteijn. Visual discomfort and visual fatigue of stereoscopic displays: a review. *Journal of Imaging Science and Technology*, 53(3):30201–1, 2009. (Cited on page 39.)

[85] Bettina Laugwitz, Theo Held, and Martin Schrepp. *Construction and evaluation of a user experience questionnaire*. Springer, 2008. (Cited on page 70.)

[86] Jeremy Laviole and Martin Hachet. PapARt: Interactive 3D graphics and multi-touch augmented paper for artistic creation. *2012 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 3–6, mar 2012. (Cited on pages 44 and 119.)

[87] Austin Lee, Hiroshi Chigira, Sheng Kai Tang, Kojo Acquah, and Hiroshi Ishii. Annoscape: remote collaborative review using live video overlay in shared 3d virtual workspace. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*, pages 26–29. ACM, 2014. (Cited on page 110.)

[88] Jeong-Eom Lee, Satoshi Miyashita, Kousuke Azuma, Joo-Ho Lee, and Gwi-Tae Park. Anamorphosis projection by ubiquitous display in intelligent space. In *Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*, pages 209–217. Springer, 2009. (Cited on page 19.)

[89] Jinha Lee, Alex Olwal, Hiroshi Ishii, and Cati Boulanger. Spacetop: integrating 2d and spatial 3d interactions in a see-through desktop environment. In *CHI*, pages 189–192. ACM, 2013. (Cited on page 55.)

[90] Dietrich Lehmann, Pascal L Faber, Shisei Tei, Roberto D Pascual-Marqui, Patricia Milz, and Kieko Kochi. Reduced functional connectivity between cortical sources in five meditation traditions detected with lagged coherence using EEG tomography. *NeuroImage*, 60(2):1574–86, apr 2012. (Cited on page 84.)

[91] John J Leonard and Hugh F Durrant-Whyte. Mobile robot localization by tracking geometric beacons. *Robotics and Automation, IEEE Transactions on*, 7(3):376–382, 1991. (Cited on page 29.)

[92] Ian Li, Anind K Dey, and Jodi Forlizzi. Understanding my data, myself: supporting self-reflection with ubicomp technologies. In *Proceedings of the 13th international conference on Ubiquitous computing*, pages 405–414. ACM, 2011. (Cited on page 93.)

[93] David G Lowe. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, volume 2, pages 1150–1157. Ieee, 1999. (Cited on page 29.)

[94] E. T. A. Maas, M. R. Marner, R. T. Smith, and B. H. Thomas. Supporting freeform modelling in spatial augmented reality environments with a new deformable material. In H. Shen and R. T. Smith, editors, *Australasian User Interface Conference (AUIC 2012)*, volume 126 of *CRPIT*, pages 77–86, Melbourne, Australia, 2012. ACS. (Cited on page 29.)

[95] Wendy E Mackay. Augmented reality: linking real and virtual worlds: a new paradigm for interacting with computers. In *Proceedings of the working conference on Advanced visual interfaces*, pages 13–21. ACM, 1998. (Cited on pages 18 and 19.)

[96] I. Scott MacKenzie. Movement time prediction in human-computer interfaces. pages 483–493, 1992. (Cited on pages 44, 49, and 51.)

[97] Michael R Marner and Bruce H Thomas. Augmented foam sculpting for capturing 3d models. In *3DUI*, pages 63–70. IEEE, 2010. (Cited on page 54.)

[98] Michael R. Marner, Bruce H. Thomas, and Christian Sandor. Physical-virtual tools for spatial augmented reality user interfaces. In *ISMAR*, pages 205–206. IEEE, oct 2009. (Cited on pages 45, 54, and 56.)

[99] Diego Martinez Plasencia, Edward Joyce, and Sriram Subramanian. Mistable: reach-through personal screens for tabletops. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 3493–3502. ACM, 2014. (Cited on page 39.)

[100] David McCandless. *Information is Beautiful*. HarperCollins, 2010. ISBN 978-0007294664. (Cited on page 94.)

[101] Rollin McCraty, Mike Atkinson, Dana Tomasino, and Raymond Trevor Bradley. The coherent heart: Heart-brain interactions, psychophysiological coherence, and the emergence of system-wide order. *Integral Review*, 5(2):10–115, 2009. (Cited on page 106.)

[102] Sebastián Mealla, Aleksander Väljamäe, Mathieu Bosi, and Sergi Jordà. Listening to your brain: Implicit interaction in collaborative music performances. In *The International Conference on New Interfaces for Musical Expression (NIME)*, 2011. (Cited on page 96.)

[103] Jonathan Mercier-Ganady, Fabien Lotte, Emilie Loup-Escande, Maud Marchal, and Anatole Lécuyer. The mind-mirror: See your brain in action in your head using eeg and augmented reality. In *Virtual Reality (VR), 2014 iEEE*, pages 33–38. IEEE, 2014. (Cited on pages 78 and 79.)

[104] Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994. (Cited on pages 17 and 18.)

[105] Gary Garcia Molina, Tsvetomira Tsoneva, and Anton Nijholt. Emotional brain-computer interfaces. *Int J Auton Adapt Commun Syst*, 6(1):9, 2009. (Cited on page 101.)

[106] Neema Moraveji, Ben Olson, Truc Nguyen, Mahmoud Saadat, Yaser Khalighi, Roy Pea, and Jeffrey Heer. Peripheral paced respiration: influencing user physiology during information work. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 423–428. ACM, 2011. (Cited on page 122.)

[107] Neema Moraveji, Athman Adiseshan, and Takehiro Hagiwara. Breathtray: augmenting respiration self-regulation without cognitive deficit. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 2405–2410. ACM, 2012. (Cited on page 122.)

[108] Daniel Moreno and Gabriel Taubin. Simple, accurate, and robust projector-camera calibration. In *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on*, pages 464–471. IEEE, 2012. (Cited on pages 26 and 32.)

[109] Tim Mullen, Christian Kothe, Yu Mike Chi, Alejandro Ojeda, Trevor Kerth, Scott Makeig, Gert Cauwenberghs, and Jung Tzyy-Ping. Real-time modeling and 3D visualization of source dynamics and connectivity using wearable EEG. 2013:2184–2187, 2013. (Cited on page 77.)

[110] Miguel A. Nacenta, Samer Sallam, Bernard Champoux, Sriram Subramanian, and Carl Gutwin. Perspective cursor: Perspective-based interaction for multi-display environments. In *CHI*, pages 289–298. ACM, 2006. (Cited on pages 45 and 52.)

[111] Richard A Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J Davison, Pushmeet Kohi, Jamie Shotton, Steve Hodges, and Andrew Fitzgibbon. Kinectfusion: Real-time dense surface mapping and tracking. In *ISMAR*, pages 127–136. IEEE, 2011. (Cited on pages 26 and 60.)

[112] Ernst Niedermeyer and FH Lopes da Silva. *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005. (Cited on pages 75, 76, 77, and 84.)

[113] Richard E. Nisbett and Timothy DeCamp Wilson. Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3):231–260, 1977. (Cited on pages 96 and 98.)

[114] Donald A Norman. *The psychology of everyday things*. Basic books, 1988. (Cited on page 4.)

[115] Donald A. Norman. Affordance, conventions, and design. *interactions*, 6(3):38–43, may 1999. (Cited on page 4.)

[116] Leyla Norooz, Matthew L Mauriello, Anita Jorgensen, Brenna McNally, and Jon E Froehlich. Bodyvis: A new approach to body learning through wearable sensing and visualization. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1025–1034. ACM, 2015. (Cited on pages 79 and 95.)

[117] B S Oken, M C Salinsky, and S M Elsas. Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clin Neurophysiol*, 117(9), sep 2006. (Cited on page 101.)

[118] Mai Otsuki, Kenji Sugihara, Asako Kimura, Fumihisa Shibata, and Hideyuki Tamura. Mai painting brush: An interactive device that realizes the feeling of real painting. In *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, pages 97–100, New York, NY, USA, 2010. ACM. (Cited on page 54.)

[119] Ana Paiva, João Dias, Daniel Sobral, Ruth Aylett, Sarah Woods, Lynne Hall, and Carsten Zoll. Learning by Feeling: Evoking Empathy with Synthetic Characters. *Applied Artificial Intelligence*, 19(3-4):235–266, 2005. (Cited on page 91.)

[120] R.D. Pascual-Marqui. Standardized low resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find Exp Clin Pharmacol*, 24D:5–12, 2002. (Cited on page 81.)

[121] Robert Patterson. Human factors of 3-D displays. *Journal of the Society for Information Display*, 15(11):861, 2007. (Cited on page 39.)

[122] G. Pfurtscheller and F. H. Lopes da Silva. Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin Neurophysiol*, 110(11):1842–1857, 1999. (Cited on page 84.)

[123] Rosalind W. Picard. Affective computing. Technical Report 321, MIT Media Laboratory, 1995. (Cited on pages 7, 9, and 94.)

[124] Ben Piper, Carlo Ratti, and Hiroshi Ishii. Illuminating clay: a 3-d tangible interface for landscape analysis. In *CHI*, pages 355–362. ACM, 2002. (Cited on page 54.)

[125] Ming-Zher Poh, Nicholas Swenson, and Rosalind W. Picard. A wearable sensor for unobtrusive, long-term assesment of electrodermal activity. *IEEE Trans. Biomed. Eng*, 57(5):1243–1252, 2010. (Cited on page 100.)

[126] Mirko Raca and Pierre Dillenbourg. System for Assessing Classroom Attention. In *LAK '13*, pages 1–5, 2013. (Cited on page 110.)

[127] Ramesh Raskar, Greg Welch, Matt Cutts, Adam Lake, Lev Stesin, and Henry Fuchs. The office of the future: A unified approach to image-based modeling and spatially immersive displays. In *I3D*, pages 179–188. ACM, 1998. (Cited on pages 11, 19, 20, 54, and 55.)

[128] Ramesh Raskar, Greg Welch, and Henry Fuchs. Spatially augmented reality. In *IWAR*, pages 11–20. Citeseer, 1998. (Cited on pages 11, 19, and 38.)

[129] Ramesh Raskar, Greg Welch, Kok-Lim Low, and Deepak Bandyopadhyay. Shader lamps: Animating real objects with image-based illumination. In StevenJ. Gortler and Karol Myszkowski, editors, *Eurographics*, pages 89–102. Springer Vienna, 2001. (Cited on pages 11 and 22.)

[130] Jun Rekimoto and Masanori Saitoh. Augmented surfaces: a spatially continuous work space for hybrid computing environments. In *CHI*, pages 378–385. ACM, 1999. (Cited on pages 46, 54, 55, 57, and 58.)

[131] Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, O. Bertrand, and A. Lécuyer. OpenViBE: An Open-Source Software Platform to Design, Test and Use Brain-Computer Interfaces in Real and Virtual Environments. *Presence-Teleop Virt*, 19(1):35–53, 2010. (Cited on pages 77 and 85.)

[132] Brett Ridel, Patrick Reuter, Jérémy Laviole, Nicolas Mellado, Nadine Couture, and Xavier Granier. The revealing flashlight:

Interactive spatial augmented reality for detail exploration of cultural heritage artifacts. *Journal on Computing and Cultural Heritage (JOCCH)*, 7(2):6, 2014. (Cited on page 38.)

[133] Yvonne Rogers. Moving on from weiser's vision of calm computing: Engaging ubicomp experiences. In *UbiComp 2006: Ubiquitous Computing*, pages 404–421. Springer, 2006. (Cited on pages 8 and 123.)

[134] David Rose. *Enchanted Objects: Design, Human Desire, and the Internet of Things*. Simon and Schuster, 2014. (Cited on pages 105 and 123.)

[135] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: an efficient alternative to sift or surf. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 2564–2571. IEEE, 2011. (Cited on page 29.)

[136] Daniel Saakes, Masahiko Inami, Takeo Igarashi, Naoya Koizumi, and Ramesh Raskar. Shader printer. In *ACM SIGGRAPH 2012 Emerging Technologies*, page 18. ACM, 2012. (Cited on page 69.)

[137] Munehiko Sato, Ivan Poupyrev, and Chris Harrison. Touché: Enhancing touch interaction on humans, screens, liquids, and everyday objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 483–492, New York, NY, USA, 2012. ACM. (Cited on page 44.)

[138] Leila Schemali and Elmar Eisemann. Design and evaluation of mouse cursors in a stereoscopic desktop environment. In *3D User Interfaces (3DUI), 2014 IEEE Symposium on*, pages 67–70. IEEE, 2014. (Cited on page 44.)

[139] Michael Schmitz. *Tangible interaction with anthropomorphic smart objects in instrumented environments*. PhD thesis, Universität des Saarlandes, 2011. (Cited on page 97.)

[140] F Schober, R Schellenberg, and W Dimpfel. Reflection of mental exercise in the dynamic quantitative topographical EEG. *Neuropsychobiology*, 31(2), 1995. (Cited on page 101.)

[141] Sue Ann Seah, Diego Martinez Plasencia, Peter D Bennett, Abhijit Karnik, Vlad Stefan Otrocol, Jarrod Knibbe, Andy Cockburn, and Sriram Subramanian. Sensabubble: a chrono-sensory mid-air display of sight and smell. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 2863–2872. ACM, 2014. (Cited on page 39.)

[142] Orit Shaer and Eva Hornecker. Tangible User Interfaces: Past, Present, and Future Directions. *Foundations and Trends*

in Human-Computer Interaction, 3(1-2):1–137, 2009. (Cited on page 54.)

[143] Daniel J Simons and Christopher F Chabris. What people believe about how memory works: A representative survey of the US population. *PloS one*, 6(8):e22757, 2011. (Cited on page 78.)

[144] Joan Sol Roo, Renaud Gervais, and Martin Hachet. Inner Garden: an Augmented Sandbox Designed for Self-Reflection. In *Proceedings of the 10th International Conference on Tangible, Embedded and Embodied Interaction*, TEI '16, Eindhoven, Netherlands, feb 2016. ACM. Work-in-Progress.

[145] Martin Spindler, Wolfgang Büschel, and Raimund Dachselt. Use your head: tangible windows for 3d information spaces in a tabletop environment. In *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces*, pages 245–254. ACM, 2012. (Cited on page 55.)

[146] Brett Stevens, Jennifer Jerrams-Smith, David Heathcote, and David Callear. Putting the virtual into reality: assessing object-presence with projection-augmented models. *Presence: Teleoperators and Virtual Environments*, 11(1):79–92, 2002. (Cited on page 37.)

[147] Arkadiusz Stopczynski, Carsten Stahlhut, Jakob Eg Larsen, Michael Kai Petersen, and Lars Kai Hansen. The Smartphone Brain Scanner: A Portable Real-Time Neuroimaging System. *PloS one*, 9(2):e86733, 2014. (Cited on page 77.)

[148] Ivan E Sutherland. The ultimate display. *Multimedia: From Wagner to virtual reality*, 1965. (Cited on pages 10 and 17.)

[149] Ivan E. Sutherland. A head-mounted three dimensional display. *Proceedings of the December 9-11, 1968, fall joint computer conference, part I on - AFIPS '68 (Fall, part I)*, page 757, 1968. (Cited on page 17.)

[150] François Tadel, Sylvain Baillet, John C Mosher, Dimitrios Pantazis, and Richard M Leahy. Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput Intell Neurosci*, 2011: 879716, jan 2011. (Cited on page 81.)

[151] Desney S Tan and Mary Czerwinski. Effects of visual separation and physical discontinuities when distributing information across multiple displays. In *INTERACT*, volume 3, pages 252–255, 2003. (Cited on page 59.)

[152] DS Tan and A Nijholt. *Brain-Computer Interaction: Applying our Minds to Human-Computer Interaction*. Springer-Verlag: London., 2010. (Cited on page 76.)

[153] Robert J. Teather and Wolfgang Stuerzlinger. Pointing at 3D targets in a stereo head-tracked virtual environment. *2011 IEEE Symposium on 3D User Interfaces (3DUI)*, (1):87–94, mar 2011. (Cited on page 46.)

[154] Robert J. Teather and Wolfgang Stuerzlinger. Pointing at 3d target projections with one-eyed and stereo cursors. In *CHI*, pages 159–168. ACM, 2013. (Cited on pages 44 and 46.)

[155] Emanuele Trucco and Alessandro Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1998. ISBN 0132611082. (Cited on page 127.)

[156] Sherry Turkle. *Alone together: Why we expect more from technology and less from each other*. Basic books, 2012. (Cited on page 7.)

[157] Brygg Ullmer and Hiroshi Ishii. The metadesk: models and prototypes for tangible user interfaces. In *Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 223–232. ACM, 1997. (Cited on page 117.)

[158] John Underkoffler and Hiroshi Ishii. Illuminating light: An optical design tool with a luminous-tangible interface. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 542–549. ACM Press/Addison-Wesley Publishing Co., 1998. (Cited on page 22.)

[159] John Underkoffler and Hiroshi Ishii. Urp: a luminous-tangible workbench for urban planning and design. In *CHI*, pages 386–393. ACM, 1999. (Cited on page 22.)

[160] Seçil Uğur. *Wearing Embodied Emotions: A Practice Based Design Research on Wearable Technology*. Springer Milan, 2013. ISBN 978-88-470-5246-8. (Cited on page 95.)

[161] S Valins. Cognitive effects of false heart-rate feedback. *Journal of personality and social psychology*, 4(4):400–8, oct 1966. (Cited on page 96.)

[162] Peter Vandoren, Tom Van Laerhoven, Luc Claesen, Johannes Taelman, Chris Raymaekers, and Frank Van Reeth. Intupaint: Bridging the gap between physical and digital painting. In *Horizontal Interactive Human Computer Systems, 2008. TABLETOP 2008. 3rd IEEE International Workshop on*, pages 65–72. IEEE, 2008. (Cited on page 54.)

[163] Jouke C Verlinden, A De Smit, Aernout WJ Peeters, and Martijn H van Gelderen. Development of a flexible augmented prototyping system. *WSCG*, 11, 2003. (Cited on page 55.)

[164] Roel Vertegaal. The (re)usability of everyday computational things: why industrial design will be the new interaction design. *interactions*, 18(1):38–41, jan 2011. (Cited on page 10.)

[165] B. Victor. A brief rant on the future of interaction design, Nov 2011. URL http://worrydream.com/ABriefRantOnTheFutureOfInteractionDesign/. (Cited on page 3.)

[166] Bret Victor. Up and down the ladder of abstraction, 2011. URL http://worrydream.com/#!2/LadderOfAbstraction. (Cited on page 110.)

[167] Bret Victor. Humane representation of thought: a trail map for the 21st century. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, pages 699–699. ACM, 2014. (Cited on pages 3, 9, and 35.)

[168] Thomas Vincent, Laurence Nigay, and Takeshi Kurata. Precise pointing techniques for handheld augmented reality. In *Human-Computer Interaction–INTERACT 2013*, pages 122–139. Springer, 2013. (Cited on pages 43 and 118.)

[169] Daniel Vogel and Géry Casiez. Conté: multimodal input inspired by an artist's crayon. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 357–366. ACM, 2011. (Cited on page 54.)

[170] S. Vosniadou, C. Ioannides, A. Dimitrakopoulou, and E. Papademetriou. Designing learning environments to promote conceptual change in science. *Learn Instr*, (11):381–419, 2001. (Cited on pages 76 and 78.)

[171] Stella Vosniadou. Knowledge Acquisition and Conceptual Change. *Applied Psychology*, 41(4):347–357, 1992. (Cited on pages 78 and 91.)

[172] Daniel Wagner and Dieter Schmalstieg. Artoolkitplus for pose tracking on mobile devices. In *Proceedings of 12th Computer Vision Winter Workshop (CVWW'07)*, pages 139–146, 2007. (Cited on pages 28, 29, and 85.)

[173] Wouter Walmink, Danielle Wilde, and Florian'Floyd' Mueller. Displaying heart rate data on a bicycle helmet to support social exertion experiences. In *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction*, pages 97–104. ACM, 2014. (Cited on pages 79 and 110.)

[174] Colin Ware. *Information visualization: perception for design*. Elsevier, 2012. (Cited on page 30.)

[175] Colin Ware, Kevin Arthur, and Kellogg S Booth. Fish tank virtual reality. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pages 37–42. ACM, 1993. (Cited on page 33.)

[176] Mark Weiser. The computer for the 21st century. *Scientific american*, 265(3):94–104, 1991. (Cited on page 8.)

[177] Mark Weiser. Some computer science issues in ubiquitous computing. *Communications of the ACM*, 36(7):75–84, 1993. (Cited on page 8.)

[178] Mark Weiser and John Seely Brown. Designing calm technology, 1995. *PowerGrid Journal*, 1(1):75–85, 1996. (Cited on pages 8 and 123.)

[179] Jerry Wellington. Formal and informal learning in science: The role of the interactive science centres. *Physics Education*, 25(5): 247, 1990. (Cited on page 78.)

[180] Pierre Wellner. Interacting with paper on the digitaldesk. *Communications of the ACM*, 36(7):87–96, 1993. (Cited on page 19.)

[181] Pierre Wellner, Wendy Mackay, and Rich Gold. Back to the real world. *Communications of the ACM*, 36(7):24–26, 1993. (Cited on pages 8, 9, and 18.)

[182] Michele A Williams, Asta Roseway, Chris O'Dowd, Mary Czerwinski, and Meredith Ringel Morris. Swarm: An actuated wearable for mediating affect. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, pages 293–300. ACM, 2015. (Cited on pages 79 and 95.)

[183] Karl Willis, Eric Brockmeyer, Scott Hudson, and Ivan Poupyrev. Printed optics: 3d printing of embedded optical elements for interactive devices. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*, pages 589–598. ACM, 2012. (Cited on page 26.)

[184] Karl DD Willis, Takaaki Shiratori, and Moshe Mahler. Hideout: mobile projector interaction with tangible objects and surfaces. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, pages 331–338. ACM, 2013. (Cited on page 29.)

[185] Bob G Witmer and Michael J Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments*, 7(3):225–240, 1998. (Cited on page 37.)

[186] Gary Wolf. The data-driven life. *The New York Times*, 28:MM38, 2010. (Cited on page 93.)

[187] J. R. Wolpaw and E. W. Wolpaw. *BCI: Principles and Practice*. Oxford University Press, 2012. (Cited on pages 9, 76, 77, and 92.)

[188] Jonathan R Wolpaw, Niels Birbaumer, William J Heetderks, Dennis J McFarland, P Hunter Peckham, Gerwin Schalk, Emanuel Donchin, Louis A Quatrano, Charles J Robinson, Theresa M Vaughan, et al. Brain-computer interface technology: a review of the first international meeting. *IEEE transactions on rehabilitation engineering*, 8(2):164–173, 2000. (Cited on page 9.)

[189] Kanit Wongsuphasawat, Alex Gamburg, and Neema Moraveji. You can't force calm: designing and evaluating respiratory regulating interfaces for calming technology. In *Adjunct proceedings of the 25th annual ACM symposium on User interface software and technology*, pages 69–70. ACM, 2012. (Cited on page 108.)

[190] Robert Xiao, Miguel A Nacenta, Regan L Mandryk, Andy Cockburn, and Carl Gutwin. Ubiquitous cursor: a comparison of direct and indirect pointing feedback in multi-display environments. In *GI*, pages 135–142. Canadian Human-Computer Communications Society, 2011. (Cited on page 45.)

[191] Robert Xiao, Chris Harrison, and Scott E Hudson. Worldkit: rapid and easy creation of ad-hoc interactive applications on everyday surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 879–888. ACM, 2013. (Cited on page 55.)

[192] Thorsten O Zander and Christian Kothe. Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. *J. Neural. Eng*, 8 (2), apr 2011. (Cited on page 94.)

[193] Shumin Zhai. Characterizing computer input with Fitts' law parameters – the information and non-information aspects of pointing. *Int. J. Hum.-Comput. St.*, 61(December):1–17, 2004. (Cited on pages 50 and 51.)

[194] Zhengyou Zhang. A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(11):1330–1334, 2000. (Cited on page 32.)