Department of Informatics

University of Fribourg (Switzerland)

A FRAMEWORK FOR ABSTRACTING, DESIGNING AND BUILDING TANGIBLE GESTURE INTERACTIVE SYSTEMS

THESIS

presented to the Faculty of Science, University of Fribourg (Switzerland)

in consideration for the award of the academic grade of

Doctor Scientiarum Informaticarum

by

Leonardo Angelini

from

Terni (Italy)

Thesis No: 2003

UniPrint, Fribourg

2017

Accepted by the Faculty of Science of the University of Fribourg (Switzerland) upon the recommendation of

- Prof. Denis Lalanne, University of Fribourg, Switzerland (Thesis Director)
- Prof. Elena Mugellini, University of Applied Sciences and Arts Western Switzerland, Switzerland (Thesis Co-Advisor)
- Prof. Philippe Cudré-Mauroux, University of Fribourg, Switzerland (Jury President)
- Prof. Patrizia Marti, Eindhoven University of Technology (TU/e), the Netherlands, and University of Siena, Italy (External Expert)
- Prof. Lorenzo Sommaruga, University of Applied Sciences and Arts of Southern Switzerland (SUPSI), Switzerland (External Expert)

Fribourg, January 27th, 2017

Thesis Director	Faculty Dean
Prof. Denis Lalanne	Prof. Christian Bochet

Acknowledgments

This thesis could not have been accomplished without the help of many people. I owe special thanks to my supervisors, Denis Lalanne from University of Fribourg and Elena Mugellini and Omar Abou Khaled from HumanTech Institute of HES-SO, who supported me throughout this entire thesis and, most important, who allowed me to learn how to stand on my own feet. This was a long but really enriching experience that changed my life and myself, and not only from a professional point of view. For this, I have to thank all the distant relatives, friends and, in particular, my parents, who, more or less quietly, tolerated my physical (and sometimes social) distance. Luckily, I found other friends who made it worth to live in Switzerland, beyond the professional advantages; I would not mention all of you here, but if you feel touched by this message, it surely means you are one of them. A special thank also for Antonio, for all the good time spent walking across Switzerland and beyond. Among all, I need to praise also my flatmates for supporting (and standing) me in the most stressful periods of my thesis. Redina and Manuel in particular, who were present during the thesis writing and fed me with the best delicacies that can be found in Fribourg. I gratefully thank my colleagues, especially Francesco, Stefano and Simon for all the useful discussions we had, for the time enjoyed together while collaborating on common projects and for holding me and catching me up whenever I was falling down (literally). A special mention for Maurizio, who belongs to all the formers, as a friend, colleague, flatmate, et cetera. To conclude, I owe much to Switzerland in general, allowing me to conduct research in an efficient and rewarding country and to the Hasler Foundation in particular, which funded this thesis under the “Living in smart environment” program. However, I think I owe something more to Italy, first, for all the resources spent for forming me before this PhD, and second, as an eternal source of inspiration thanks to all people who live and lived there and that shaped my thoughts and my mind. When I was a child, my aunt Rita was used to sing me a song that says: “if I were Leonardo (da Vinci), how many things I could invent…!”. I am not sure that I invented something that is actually worth my name, but I think I should still be thankful to Leonardo da Vinci for inspiring me as an engineer, inventor and artist, and for drawing the Vitruvian man, source of inspiration of the HumanTech institute. Indeed, this popular image remembers us that humans are at the center of our research and technology is only a support for reaching the human well-being.

Ringraziamenti

Non avrei potuto terminare questa tesi senza l'aiuto di molte persone. Devo ringraziare in particolare i miei supervisori, Denis Lalanne dell'Università di Fribourg ed Elena Mugellini e Omar Abou Khaled dell'Istituto HumanTech della HES-SO, che mi hanno supportato durante tutta la durata di questa tesi, in particolar modo durante le fasi più difficili, permettendomi di imparare a camminare con le mie proprie gambe. Questa tesi è stato un viaggio lungo, ma allo stesso tempo è stata anche un'esperienza che mi ha arricchito molto, che ha cambiato la mia vita e me stesso, e non solo da un punto di vista meramente professionale. Per questo, devo ringraziare tutti i miei parenti e amici lontani, ed in particolar modo i miei genitori, che hanno saputo tollerare con pazienza (chi più, chi meno) la mia distanza fisica e a volte anche sociale. Per fortuna ho potuto trovare altri amici che hanno permesso di rendere piacevole la mia vita in Svizzera, al di là dei vantaggi professionali; non potrò menzionarvi tutti qui, ma se vi sentirete toccati da queste parole, vuol dire sicuramente che site uno di loro. Un ringraziamento speciale per Antonio, per tutti i bei momenti passati camminando per la Svizzera e oltre. Tra gli altri, devo lodare i miei coinquilini per avermi supportato (e sopportato) duranti le fasi più stressanti della tesi. In particolare, Redina e Manuel, che erano presenti durante la fase di scrittura ed ebbero l'accortezza di nutrirmi con le migliori leccornie che si possano trovare a Fribourg. Ringrazio calorosamente tutti i miei colleghi, in particolare, Francesco, Stefano e Simon per tutte le discussioni fatte, per il tempo passato insieme a collaborare in progetti comuni e per sorreggermi e riprendermi tutte le volte che sono caduto. Menzione speciale per Maurizio, che fa parte di tutte le precedenti categorie, come amico, collega, coinquilino, et cetera. Per concludere, devo molto alla Svizzera in generale, che mi ha permesso di portare avanti ricerche in un paese efficiente e gratificante, ed in particolare alla Fondazione Hasler, che ha finanziato questa tesi nel contesto del programma "Vivere negli ambienti intelligenti". Nondimeno, penso di dovere qualcosa in più all'Italia; prima di tutto, per tutte le risorse impiegate nella mia formazione prima di questo PhD, e in secondo luogo, per essere fonte eterna di ispirazione, grazie a tutte le persone che vivono ed hanno vissuto in Italia e che hanno formato la mia mente e il mio modo di pensare. Quando ero bambino, mia zia Rita era solita cantarmi una canzone che fa: "se fossi Leonardo (da Vinci) quante cose inventerei…!". Non sono sicuro di aver inventato qualcosa che sia davvero degno del mio nome, ma penso ad ogni modo di essere molto grato a Leonardo da Vinci per avermi ispirato come ingegnere, inventore e artista, e soprattutto per aver disegnato l'Uomo Vitruviano, fonte di ispirazione per l'Istituto HumanTech. Infatti, questa popolare immagine ci deve ricordare che al centro della nostra ricerca c'è l'uomo, e che la tecnologia funge soltanto da supporto per il benessere umano.

Abstract

This thesis discusses tangible gesture interaction, a novel paradigm for interacting with computer that blends concepts from the more popular fields of tangible interaction and gesture interaction. Taking advantage of the human innate abilities to manipulate physical objects and to communicate through gestures, tangible gesture interaction is particularly interesting for interacting in smart environments, bringing the interaction with computer beyond the screen, back to the real world. Since tangible gesture interaction is a relatively new field of research, this thesis presents a conceptual framework that aims at supporting future work in this field. The Tangible Gesture Interaction Framework provides support on three levels. First, it helps reflecting from a theoretical point of view on the different types of tangible gestures that can be designed, physically, through a taxonomy based on three components (move, hold and touch) and additional attributes, and semantically, through a taxonomy of the semantic constructs that can be used to associate meaning to tangible gestures. Second, it helps conceiving new tangible gesture interactive systems and designing new interactions based on gestures with objects, through dedicated guidelines for tangible gesture definition and common practices for different application domains. Third, it helps building new tangible gesture interactive systems supporting the choice between four different technological approaches (embedded and embodied, wearable, environmental or hybrid) and providing general guidance for the different approaches. As an application of this framework, this thesis presents also seven tangible gesture interactive systems for three different application domains, i.e., interacting with the In-Vehicle Infotainment System (IVIS) of the car, the emotional and interpersonal communication, and the interaction in a smart home. For the first application domain, four different systems that use gestures on the steering wheel as interaction means with the IVIS have been designed, developed and evaluated. For the second application domain, an anthropomorphic lamp able to recognize gestures that humans typically perform for interpersonal communication has been conceived and developed. A second system, based on smart t-shirts, recognizes when two people hug and reward the gesture with an exchange of digital information. Finally, a smart watch for recognizing gestures performed with objects held in the hand in the context of the smart home has been investigated. The analysis of existing systems found in literature and of the system developed during this thesis shows that the framework has a good descriptive and evaluative power. The applications developed during this thesis show that the proposed framework has also a good generative power.

Riassunto

Questa tesi discute l'interazione gestuale tangibile, un nuovo paradigma per interagire con il computer che unisce i principi dei più comuni campi di studio dell'interazione tangibile e dell'interazione gestuale. Sfruttando le abilità innate dell'uomo di manipolare oggetti fisici e di comunicare con i gesti, l'interazione gestuale tangibile si rivela particolarmente interessante per interagire negli ambienti intelligenti, riportando l'attenzione sul nostro mondo reale, al di là dello schermo dei computer o degli smartphone. Poiché l'interazione gestuale tangibile è un campo di studio relativamente recente, questa tesi presenta un framework (quadro teorico) che ha lo scopo di assistere lavori futuri in questo campo. Il Framework per l'Interazione Gestuale Tangibile fornisce supporto su tre livelli. Per prima cosa, aiuta a riflettere da un punto di vista teorico sui diversi tipi di gesti tangibili che possono essere eseguiti fisicamente, grazie a una tassonomia basata su tre componenti (muovere, tenere, toccare) e attributi addizionali, e che possono essere concepiti semanticamente, grazie a una tassonomia di tutti i costrutti semantici che permettono di associare dei significati ai gesti tangibili. In secondo luogo, il framework proposto aiuta a concepire nuovi sistemi interattivi basati su gesti tangibili e a ideare nuove interazioni basate su gesti con gli oggetti, attraverso linee guida per la definizione di gesti tangibili e una selezione delle migliore pratiche per i differenti campi di applicazione. Infine, il framework aiuta a implementare nuovi sistemi interattivi basati su gesti tangibili, permettendo di scegliere tra quattro differenti approcci tecnologici (incarnato e integrato negli oggetti, indossabile, distribuito nell'ambiente, o ibrido) e fornendo una guida generale per la scelta tra questi differenti approcci. Come applicazione di questo framework, questa tesi presenta anche sette sistemi interattivi basati su gesti tangibili, realizzati per tre differenti campi di applicazione: l'interazione con i sistemi di infotainment degli autoveicoli, la comunicazione interpersonale delle emozioni, e l'interazione nella casa intelligente. Per il primo campo di applicazione, sono stati progettati, sviluppati e testati quattro differenti sistemi che usano gesti tangibili effettuati sul volante come modalità di interazione con il sistema di infotainment. Per il secondo campo di applicazione, è stata concepita e sviluppata una lampada antropomorfica in grado di riconoscere i gesti tipici dell'interazione interpersonale. Per lo stesso campo di applicazione, un secondo sistema, basato su una maglietta intelligente, riconosce quando due persone si abbracciano e ricompensa questo gesto con uno scambio di informazioni digitali. Infine, per l'interazione nella casa intelligente, è stata investigata la realizzazione di uno smart watch per il riconoscimento di gesti eseguiti con oggetti tenuti nella mano. L'analisi dei sistemi interattivi esistenti basati su gesti tangibili permette di dimostrare che il framework ha un buon potere descrittivo e valutativo. Le applicazioni sviluppate durante la tesi mostrano che il framework proposto ha anche un valido potere generativo.

Table of Contents

Table of Figures

1 Introduction

“Più facilmente si contasta al principio che alla fine.”

Leonardo da Vinci, “Scritti letterari, Pensieri, 67”

This chapter presents the motivation that brought to the definition of the thesis subject and of the challenges addressed in this thesis. In particular, it presents a brief history of the evolution of the human communicative and manipulative skills, arguing that most part of current digital interactions is contributing to the loss of these skills. To this purpose, this chapter introduces tangible gestures and proposes a new paradigm for digital interactions that are closer to the real world and that can better exploit humans’ natural skills.

1.1 Back to the Real World: A Challenge for Interaction Designers

During their evolution, humans developed particular abilities that allowed them to differentiate from all the other animals. Indeed, the particular conformation of the thumb, opposable to the other digits, allows humans to grasp objects and to *manipulate* tools through complex movements. Human ancestors’ ability to firmly grasp small objects and flex the wrist facilitated the evolution of technologies for processing food, for chasing animals and, in an earlier era, for crafting new tools [2].

Another important evolution of human beings concerns the particular ability to communicate complex information. When humans developed the first small communities, there was an increasing exigence of communication. Humans developed two different communication means: *gestures* were particularly important to communicate while chasing animals, allowing a silent communication that would avoid being perceived by animals; at the same time, to satisfy the increasing necessity to communicate information, humans developed articulated languages, which allowed continuous *voice* communication even while performing manual activities. Some theories suggest that first spoken languages originated from gestures [72] and then evolved together in the different cultures. Several studies demonstrated that still nowadays humans use the same cognitive processes for the generation of speech and gestures [164].

This short introduction about the origins of human-to-human communication is important to understand the history of Human-Computer Interaction and its effort towards the definition of natural interaction languages to facilitate the communication between humans and computer. Indeed, at the beginning of the digital information era, the interaction between humans and computers was reserved to few experts, who were acquainted with the low-level programming languages and command line consoles. The introduction of the Windows Icon Menu Pointer (WIMP) paradigm allowed a rapid diffusion of the Personal Computer (PC), bringing the digital information era to the mass. Everybody, after some training, could benefit from the digital services offered by the PC. Nevertheless, the PC exploits only a small part of the human sensorimotor abilities: in order to depict "how the computer sees us", O'Sullivan and Igoe used the dramatic image of a live being constituted by one hand with only one finger and provided with two ears and only one eye [150] (See Figure 1). Another limitation of WIMP interfaces is the learning phase necessary to master the mouse-pointer interaction, which is often troublesome, especially for older adults [91] .

Figure 1. "How the computer sees us" [150]

The advent of touchscreen technology allowed to facilitate the learning curve to start interacting with computer, simplifying the access to digital information also for older adults [91]. Simple gestures, such as tapping, swiping and pinching allow to interact intuitively with the Graphical User Interface (GUI) and to manipulate directly virtual objects represented in the GUI. To identify the systems that make use of facilitated interaction modalities, in particular multi-touch interfaces, the HCI community coined the term Natural User Interfaces (NUI). In other cases, the community refers to natural interaction whenever the humans' innate abilities are exploited. Since *gestures*, *voice* and *tangible* *manipulations*, as stated before, accompanied humans during their whole evolution, they are often considered all as natural interaction modalities. Nevertheless, Norman raised an important debate about which interactions can be really considered as "natural" [147]. Indeed, he argued that some gesture interfaces, even compared to graphical user interfaces (GUIs) operated by mouse and keyboard, are often not easy to learn and to understand, providing a user experience that is not perceived as natural at all.

Voice and gesture interaction with computer had been explored since the beginning of the computer era. Indeed, in 1980, Bolt's Put-That-There system already allowed to interact with digital information through gestures and voice commands [29]. Although voice could be seen as a very natural means to interact with computer, there are many situations in which voice recognition could be not effective (in noisy environment), or simply not acceptable for the user (in public or crowded environments) [63]. Many mobile and desktop operative systems integrate voice commands and vocal assistants (e.g., Apple Siri[1], Google Voice[2] and Microsoft Cortana[3]) but their usage is still limited. The 2013 sci-fi film "Her"[4] offers a critical view about the user acceptance and the possible impact in our society of such systems in a future where artificial intelligence could simulate a human live being.

Although traditional GUIs are still nowadays the most efficient, practical and socially accepted user interfaces to interact with computer in many application domains, especially in work environments, already in 1993 Wellner et al. felt the need of designing interfaces that could bring back humans to their natural environment [223]. Spending hours sitting still in front of a PC, tapping on a keyboard and slightly moving the mouse can be very economical in terms of physical effort, but these interactions exploit just a minimal percentage of the human physical abilities. To cope with this trend, Wellner et al. suggested to augment with computation our everyday environment and activities, interacting directly with the physical objects of our everyday life, instead that mediating the interaction through virtual elements presented by a computer [223]. The authors identified two already existing research fields that embraced this philosophy: ubiquitous interaction and augmented reality.

Weiser's vision of ubiquitous computing [222], with technology that is waved into the fabric of the everyday life and becomes transparent to the user, completely fits Wellner et al.'s [223] suggestion of bringing back the interaction to the real world. Nevertheless, the most successful ubiquitous technology that is nowadays seamlessly integrated into the humans' everyday life is probably the smartphone. Although the smartphone allowed moving digital interaction beyond the desktop, it still

---

[1] http://www.apple.com/ios/siri/
[2] https://www.google.com/googlevoice/about.html
[3] https://support.microsoft.com/en-us/help/17214
[4] http://www.imdb.com/title/tt1798709/

offers a limited interface in terms of the exploitation of human natural skills. The aforementioned image of O'Sullivan and Igoe [150] (Figure 1) can be used to depict also "how the smartphone sees us": as for the PC, most smartphone interactions involve only one hand, one finger, one eye and two ears. Paradoxically, while the smartphone has become part of humans' everyday routine, bringing the digital interaction beyond the desktop and embodying the ubiquitous computing philosophy, at the same time, it contributed to detach humans from the real world even more than traditional GUIs. Allowing humans to being connected to the virtual world anywhere and anytime, it often profoundly modified their behavior in social contexts. People interacting with their smartphone seem often closed in their digital bubbles, ignoring other people and things happening around them. This is a dangerous behavior, not only because it implies a decline of human social skills, but also because it detaches the human cognition and attention from the real world, which can have even fatal consequences, as the Lausanne Police recently warned with a provoking video[5]. As a matter of fact, in the ubiquitous computing era, there is an increasing need of richer interactions that exploit better human communicative and manipulative skills and that better integrate in our daily routines.

A particular research field in HCI completely embraced Wellner et al.'s vision [223] of bringing back humans to the real world: tangible interaction [94]. Tangible interaction exploits the affordances of physical objects and embodies digital interactions directly in the real world. In tangible interaction, humans can benefit of their manipulative skills to access and modify the digital information through physical objects. Moreover, the interfaces distributed in the real world offer multiple accesses to the digital information and collaboration and social exchanges are naturally facilitated. Despite of these advantages, tangible interaction is integrating slowly into the users' everyday life, influencing many product and interface designs, but with few examples of tangible user interfaces for everyday use.

In the context of the environments where we spend most of our time, the technological progress is bringing computation and digital functionalities into everyday objects. Many of these objects boast also connectivity capabilities, falling into the growing category of the Internet of Things (IoT) [16]. Nowadays, IoT objects can provide much digital information, but it can generally be accessed only through traditional graphical user interfaces, via smartphones or PCs [177]. Following the tangible interaction paradigm, designers could explore more natural means to access this information directly from the object. Nevertheless, often objects do not offer enough affordances to allow tangible manipulations for each piece of information that could be accessed. Moreover, adding new interactions to an existing physical object (or once the object is already planned for the production) and

[5] Lausanne Police in conjunction with the bfu – Swiss Council for Accident Prevention, "Making magic in traffic with a smartphone" https://www.youtube.com/watch?v=-YA8S2w5Q04

personalizing the experience for each end-user can be particularly difficult and expensive[177]. In this case, gestures can offer a richer and customizable vocabulary of interactions to communicate with physical objects that embody digital information. The next sections will discuss more in depth the possibility offered by tangible gestures to interact with objects in smart environments, i.e., the main idea that drove the research conducted during this thesis.

1.2 Tangible Gestures: Interacting with Everyday Objects

Manipulating physical objects and gesturing for communicative purposes are two skills deeply rooted in the human history and that humans learn since their infancy. In particular, in order to provide additional meaning to their discourse and to facilitate the understanding to the listener, humans often use objects while gesturing. In music, gestures, defined as movements of the body, are used as a form of artistic communication [76]. These gestures are often performed with objects that act as instruments to produce sounds, extending the abilities of the human body and increasing our communication expressivity. Gestures are commonly used to communicate with other people: in particular, contact gestures performed on the human body can convey strong emotional contents (cf. Section 3.4). Radford demonstrated how gesturing with physical objects can help explaining mathematical concepts [166]. Glenberg and Robertson argued that pointing to physical objects increases the listener ability of understanding a concept, in comparison to discussing of abstract concepts without referring to the physical world that surrounds the listener [74].

These latter findings support the theories of embodied cognition, which argue that humans construct knowledge of the external word by experiencing it through their body and their senses [216]. Different visions emerge among embodied cognition theories: in particular, those arguing that the representation and the computation (in other words, the knowledge) is distributed and embodied in the objects of the real world, and those arguing that humans need to enact and experience the world in order to build the knowledge related to the objects of the real world. This latter nuance of embodied cognition highlights the importance of experiencing the external world through our body and our senses and of manipulating and gesturing with the physical objects in order to build the mental models and the knowledge that we have of the external world.

These perspectives are particularly interesting also in the field of Human-Computer Interaction, where users, in order to interact with computers, need to understand and build mental models of the interfaces. Using objects that can be physically manipulated could help the user to understand the working principles of the interface, by simply exploring the interactive objects. At the same time, designers could exploit form-giving to create objects that appeal to the user senses and motor skills [58].

The design space of gesture interaction with objects is very large: several different objects from our everyday world can be used to interact with the computer, or new ad-hoc objects can be created; a plethora of different gesture types, either with the hands or with other parts of the body, can be imagined to interact with or through these objects. Hoven and Mazalek recently formalized the emergence of a new research field in HCI, i.e., Tangible Gesture Interaction (TGI), which aims at exploring gestures with object as a rich and meaningful interaction means with computer [213]. Hoven and Mazalek individuated several examples of tangible gesture interactive systems and several application domains that can benefit of tangible gesture interaction [213]. However, they also suggested the need of further investigations to explore the design space of tangible gesture interaction. Indeed, when I started investigating this new topic at the beginning of my thesis, I found few specific previous work about tangible gesture interaction theory, although I found a large amount of literature in the field of tangible interaction and gesture interaction and several interactive systems using gestures with objects. Because of the lack of formalization of such a broad design space, I felt necessary the definition of a framework that could help future investigations in this relatively new field of research.

1.3 Thesis Context

This thesis has been carried out as a joint collaboration between the University of Fribourg and the College of Engineering and Architecture of Fribourg, member of the University of Applied Sciences and Arts Western Switzerland. This thesis is part of the project “Living in smart environments: Natural and Economic gesture-based HCI” proposed by the DIVA group[6] of the University of Fribourg and the HumanTech Institute[7] (former MISG group) of the University of Applied Sciences and Arts Western Switzerland. The project received a 4 year grant (3 year grant plus a one-year extension) from the Hasler Foundation[8], a private Swiss foundation that “promotes information and communications technology (ICT) for the well-being and benefit of Switzerland as an intellectual and industrial centre.” In particular, the aim of this thesis in the context of the project was to investigate a novel natural and economical gesture language to interact in smart environments.

[6] http://diuf.unifr.ch/main/diva/
[7] http://humantech.institute
[8] http://www.haslerstiftung.ch/

The DIVA group and the HumanTech Institute had ongoing projects for the development of smart environments (a Smart Meeting Room and a Smart Living Room in the two respective institutions). Several other theses have been carried out in the domain of gesture interaction by the DIVA group and the HumanTech Institute during the same period of my thesis. This contributed to a rich research environment, where several internal collaborations were carried out on common interest projects. Other collaborations were carried out with external academic partners, in particular with the FBK research institute of Trento, Italy, and with the Psychology department of the University of Fribourg.

Finally, the previous work of Elise van den Hoven and Ali Mazalek [213] profoundly influenced this thesis, since I had access to a preprint version of their article already in 2010, before its publication. This led to the decision of investigating tangible gesture interaction to define a natural language to interact in smart environments, which was the aim of the funded project. After an analysis of the existing literature and the definition of two possible applications, I began formalizing the design space of tangible gesture interaction, which led to the definition of a framework for this new research field. In a later time, Elise van den Hoven personally contributed with useful suggestions during the definition of the framework and the redaction of the article that presents most of the work done during this thesis [12]. Elise van Hoven and Ali Mazalek also contributed to the organization of a workshop on tangible gesture interaction, held during the TEI 2015 conference [13].

The work presented in this thesis would not have been possible without the contribution of these important actors and of many others that contributed to the definition of my research. Among other events that influenced profoundly my thesis, it is worth noting the participation to the French German Tangible Interaction Studio (FGTIS), organized by Nadine Couture (ESTIA Recherche) in Biarritz in 2013, where I had the opportunity to present a preliminary version of the Tangible Gesture Interaction Framework and to get useful feedback from the other participants. Thankful of this experience, I was glad to help in the organization of the second edition of the Tangible Interaction Studio, now called European Tangible Interaction Studio (ETIS), which was held in Fribourg in January 2016.

1.4 Specific Challenges Addressed in this Thesis

A specific challenge had to be addressed in this thesis for the “Living in smart environments” project, i.e., the definition of a natural and economical gesture language to interact in smart environments. The technological advances that are making everyday objects smarter and capable of interacting with the user oriented the choice of the gesture language toward the novel field of tangible gesture interaction, which aims at bringing the interaction into the everyday world, exploiting the affordances of the physical objects. Because of the novelty of this research field, the main research question was: *how to support the creation of new systems that exploit tangible gestures as interaction means with computer?* To answer to this broad research question, a Framework for tangible gesture interaction was needed. The Framework should be able to deal with different aspects:

- *How to formalize the syntax and the semantics of tangible gestures?* Gestures with objects can assume several forms, which vary according to the gesture type but also to the object shape. Moreover, different meanings can be associated to tangible gestures. A formalization is needed to describe the different types of tangible gestures that can be used for digital interaction purposes as well as the semantic constructs that allow associating meanings to tangible gestures.

- *How to support the design of tangible gesture interaction?* Although tangible gesture interaction stems from tangible interaction and gesture interaction, specific guidelines should be provided for designing tangible gesture interactive systems. These guidelines should benefit of the aforementioned formalization as a reference frame for enriching the design space of this novel interaction paradigm.

- *How to support the implementation of tangible gesture interactive systems?* Because of the various forms that tangible gestures can assume, several techniques can be used to recognize tangible gestures and several technologies can be used to build tangible gesture interactive systems. A review of these techniques and technologies taking into account the previous formalization could help designers building tangible gesture interactive systems.

- *How to evaluate a tangible gesture interactive system?* Different criteria should be defined in order to assess the quality of the interaction and the properties of a tangible gesture interactive system.

1.5 Thesis Structure

This thesis contains two main different parts, a theoretical framework for tangible gesture interaction and a practical application of the framework on different domains. Although the thesis presents the practical part as a logical consequence of the theoretical part, they have been carried out chronologically in parallel during this PhD. Several iterations were needed in order to obtain a complete framework for the formalization of tangible gesture interaction. At the same time, the investigations carried out for the practical applications helped understanding the elements that should be addressed by the framework in order to support the design and development of tangible gesture interactive systems.

The thesis is structured as follows. The first part of the thesis deals with the large background that can be found in the literature. In particular, Chapter 2 analyzes the theoretical background behind tangible gesture interaction, which stems from the two popular fields of tangible interaction and gesture interaction. Chapter 3 presents the literature review of the existing systems where the user interacts through tangible gestures, classified for the different application domains that were considered as relevant for this thesis.

The second part of the thesis presents the Tangible Gesture Interaction Framework (TGIF), which is the result of the theoretical investigations carried out during this thesis. The framework is structured in three parts, each dealing with one of the three different categories of tangible interaction frameworks evidenced by Mazalek and Hoven [137]: abstracting, designing and building. For abstracting, Chapter 4 proposes a communication model based on tangible gestures and analyzes the syntax and semantics of tangible gestures. For designing, Chapter 5 illustrates the typical design process, common practices for different applications domains and object affordances, typical feedback that can be provided to the user and the techniques for assessing the interaction quality and the system performances. Finally, for building, Chapter 6 shows the possible technological approaches for recognizing tangible gestures and it references the most common implementations for each approach. It also points to existing hardware toolkits and gesture recognition techniques for implementing tangible gesture interactive systems.

The third part of the framework presents the practical results of the thesis and an evaluation of the proposed framework. Indeed, in Chapter 7, seven tangible interactive systems developed during this thesis are presented and analyzed in depth according to the proposed framework: WheelSense, in its four versions, which investigates tangible gestures performed on the steering wheel to interact with infotainment system of the car; the ADA Lamp, which investigates gesture performed on an anthropomorphic lamp to communicate emotions at distance; Hugginess, a wearable system to recognize and encourage the hug gesture between people; finally, a Smart Watch to interact with everyday objects in the context of a smart home. Other systems developed during this thesis but less relevant for the proposed analysis will be briefly presented. Chapter 8 presents an evaluation of the Tangible Gesture Interactive Framework according to the three dimensions proposed by Beaudouin-Lafon [23]: the descriptive power, the evaluative power and the generative power of the framework.

The fourth part of the thesis includes the discussion and the conclusion of the thesis. Indeed, Chapter 9 discusses the results obtained in this thesis as well as the limitations and the future perspectives. Figure 2 summarize the thesis structure, evidencing which part of the thesis answers the different research questions listed in Section 1.4.

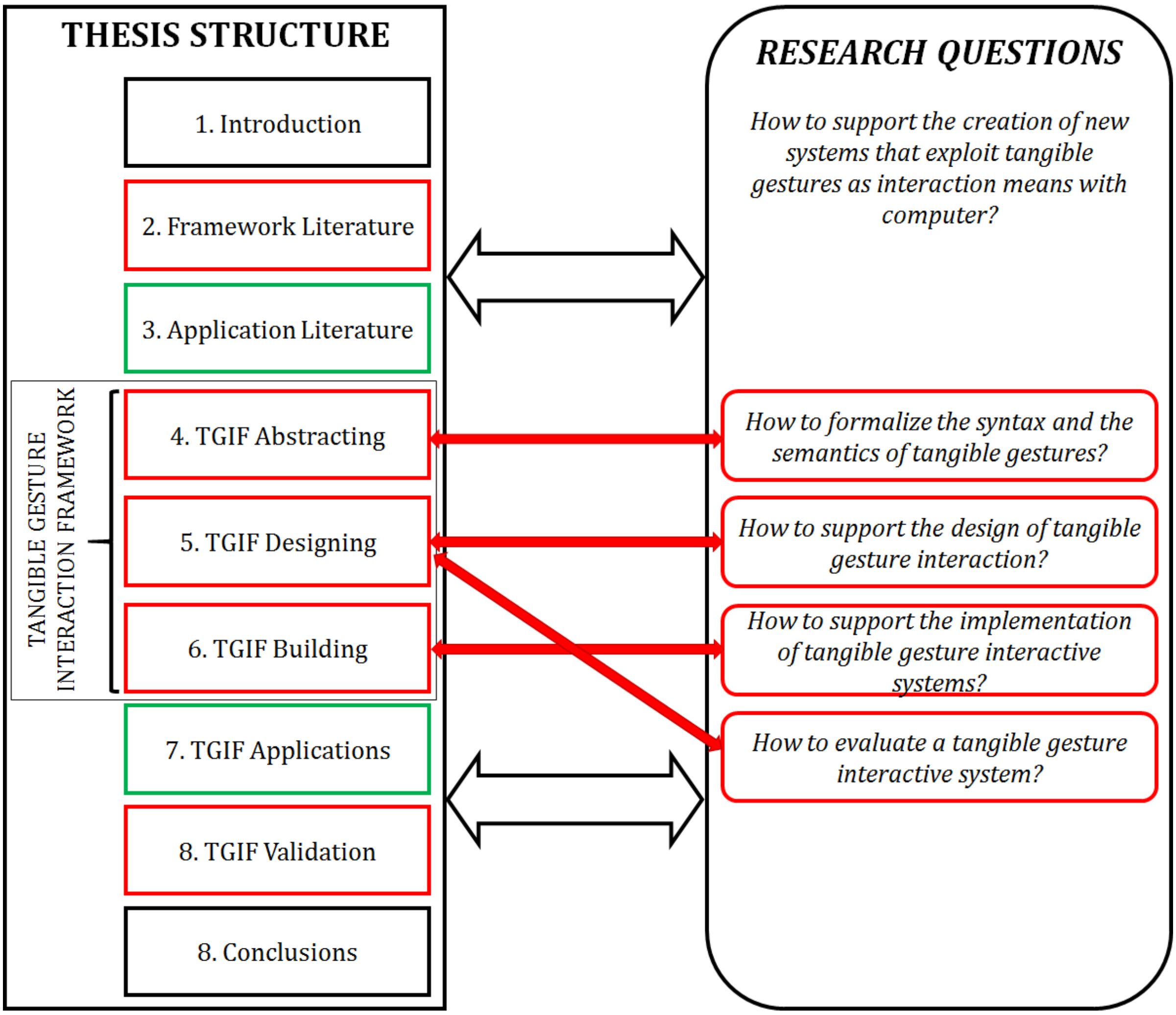

Figure 2. Thesis structure and relationship with the research questions. In red, theoretical contributions are highlighted, in green, practical contributions.

Part I – Background

2 Theories and frameworks on tangible and gesture interaction

"Sì come ogni regno in sé diviso è disfatto, così ogni ingegno diviso in diversi studi si confonde e indebolisce."
Leonardo da Vinci, "Scritti letterari, Pensieri, 27"

This chapter describes the theoretical literature linked to the main topic of this thesis, i.e., tangible gesture interaction. Since tangible gesture interaction has been recently represented at the intersection of tangible interaction and gesture interaction, bringing advances from these two fields (see Section 2.1), the related literature for tangible interaction (Section 2.2) and gesture interaction (Section 2.3) is presented. Although tangible gesture interaction has a recent history, it is worth discussing some previous works that describe from a theoretical point of view particular gestures with objects (Section 2.4). Each section frames related work into abstracting, designing and building frameworks, according to the classification of Mazalek and Hoven [137].

2.1 Definition of Tangible Gesture Interaction

In "Grasping gestures: Gesturing with physical artifacts", Hoven and Mazalek defined for the first time tangible gesture interaction as "the use of physical devices for facilitating, supporting, enhancing, or tracking *gestures* people make for digital interaction purposes. In addition, these devices meet the *tangible interaction criteria*" [213]. This article provided helpful guidance for the literature research. Indeed, Hoven and Mazalek represented tangible gesture interaction (TGI) as the intersection of tangible interaction and gesture interaction [213] (see Figure 3), bridging advantages from both approaches, i.e., the affordance of manipulating the physical world from tangible interaction and the communicative intent from gestures: even if TGI has a relatively short history, it is rooted on two well-established branches of HCI.

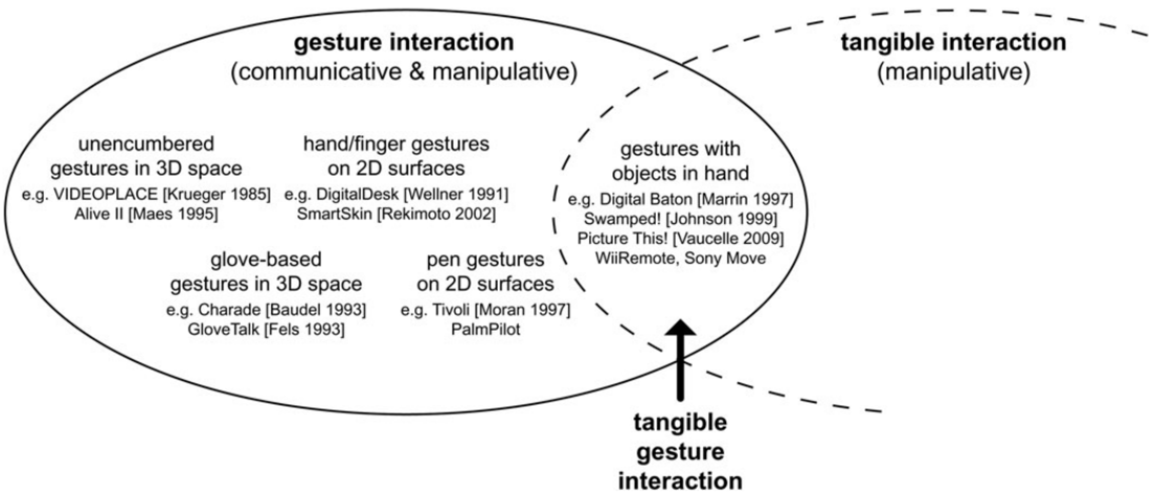

Figure 3. Representation of Tangible Gesture Interaction [213]

In particular, the previous definition of Hoven and Mazalek implicitly includes an understanding of what a *gesture* is in Human-Computer Interaction, and an understanding of the *tangible interaction criteria* cited in the definition. Hoven and Mazalek [213] referred to the Ullmer and Ishii's definition of Tangible User Interfaces [205] as criteria for tangible interaction; however, because of the rapid evolution of the tangible interaction research field, a deeper analysis was necessary to understand what tangible interaction means and to find criteria that are closer to the current definition of tangible interaction. Therefore, Chapter 2 analyzes related work in the field of tangible interaction (Section 2.2) and of gesture interaction (Section 2.3). Finally, it deals with specific researches related to gestures with objects, recently defined as tangible gesture interaction (Section 2.4).

2.2 Tangible Interaction

Tangible interaction is a broad discipline in Human-Computer Interaction: its definition is still debated in the community, because it continuously evolves and embraces new disciplines. For this reason, many researchers agree that still nowadays there is no clear and universally accepted answer to the question "What is Tangible Interaction?" (cf. [94, 185, 214]). Recently, during the IHM'15 conference, the keynote speaker Eva Hornecker was asked: "What are tangible interaction, embedded interaction and embodied interaction, *for you*?". Although Eva Hornecker answered this question in few words, she acknowledged that a complete answer would have taken hours and could have implied many debates. Debates are very popular in the field of tangible interaction (see [215] for an example) and are often generated by the different backgrounds of the heterogeneous community that participates to this research field. Tom Djajadiningrat, keynote speaker at the TEI'16 conference, raised similar concerns about the difficulty to give a definition to tangible interaction. From his point of view, tangible interaction is a generic approach to design products that exploit the human bodily skills and the meaningfulness of our everyday environment [56].

Within this evolving context, with different points of view on tangible interaction, the literature review presented in this thesis reflects a personal interpretation of what tangible interaction is *for me*. In this section, first, I briefly summarize the birth and evolution of tangible interaction, and then I analyze the most important frameworks in the field of tangible interaction. The literature review on tangible interaction presented in this thesis includes only the related work that I considered relevant for this thesis. A more extensive literature review (up to 2010) can be found in the book of Shaer and Hornecker [185].

2.2.1 Evolution of Tangible Interaction Definition

As anticipated in Section 1.1, already in the early ’90s, researchers felt the need to bring back Human-Computer Interaction to the real world [223]. Tangible interaction was born in this context and evolved then significantly over the following years. The first acknowledged system that embodied the principles of tangible interaction is the Marble Answering Machine, realized by Durrell Bishop in 1992 during his master in interaction design [28]. In the Marble Answering, voice messages are physically represented by spherical marbles and the user can play or erase messages by placing the marble in specific hollows of the machine. This is the first example of digital data embodied in a physical object. In the following years, Fitzmaurice [69] made a consistent contribution to the field, investigating during his PhD thesis the use of graspable objects to control graphical interfaces projected on tabletops. He evidenced several advantages over traditional interfaces: the objects offer multiple inputs allowing space-multiplexed interactions; moreover, the users can exploit their spatial reasoning skills and their innate ability to manipulate objects. Fitzmaurice called them Graspable User Interfaces [69], paving the way to the more popular Tangible User Interfaces (TUIs), which have been formalized in 2000 by Ullmer and Ishii [205]. Through the MCRpd interaction model, Ullmer and Ishii highlighted the importance of materializing the representation of the digital information, as well as of allowing the user to manipulate directly this information through physical objects. Ishii updated this model in 2008 [97], proposing the vision of tangible bits and introducing in the model the notion of actuation, i.e., the computer ability to modify the physical representation as additional feedback for TUIs. In 2006, Hornecker and Buur [95] presented an alternative vision to tangible interaction, going beyond the direct manipulation of typical TUIs on tabletops. Introducing full-body and spatial interactions among the typical themes and concepts that should be explored by interaction designers, they widened the definition of tangible interaction, embracing new disciplines and potential applications of tangible interaction in arts and design. Following this new trend, in 2007, the Tangible, Embedded and embodied Interaction conference was born, bringing together researchers from different backgrounds [206]. In 2012, Ishii et al. advanced the Tangible Bits vision introducing Radical Atoms and presenting a prototype of transformable matter [98]. According to their vision, in the future, the elementary components of the matter could be controlled with tangible manipulations and gestures, allowing digital interactions with transformable matter. This technology-driven vision worried part of the TEI community [215], who believes, from a design-led perspective, that a world permeated by digital matter does not offer per se any added values to the human society, while, conversely, it could rise questions about the danger of an omnipresent technology that could invade every physical aspect of our everyday environment.

This short analysis of the history of tangible interaction shows that this field is rapidly evolving and highly interdisciplinary; it is not characterized by a clear definition, but it suggests many concepts and themes that aim at improving the user experience of digital interactions in the physical world. The next subsections detail some of the most important theories and frameworks that are particular relevant for this thesis.

2.2.2 Tangible Interaction Frameworks

Several frameworks have been proposed over time to help researchers understanding the broad domain of tangible interaction and designing new tangible interfaces. According to Mazalek and Hoven [137], frameworks act as skeletal structures that other researchers could use to design and develop their work. They range from very specific to more general ones, and include guidelines, models and taxonomies. Mazalek and Hoven [137] classified tangible interaction frameworks (up to 2009) according to three types (*abstracting*, *designing* and *building*) and five facets (*experiences*, *domains*, *physicality*, *interactions* and *technologies*). Their classification is shown in Figure 4. The same classification of tangible framework types will be used in this thesis to frame the different aspects of the proposed Tangible Gesture Interaction framework. In the following subsections, the most important frameworks for each type will be presented.

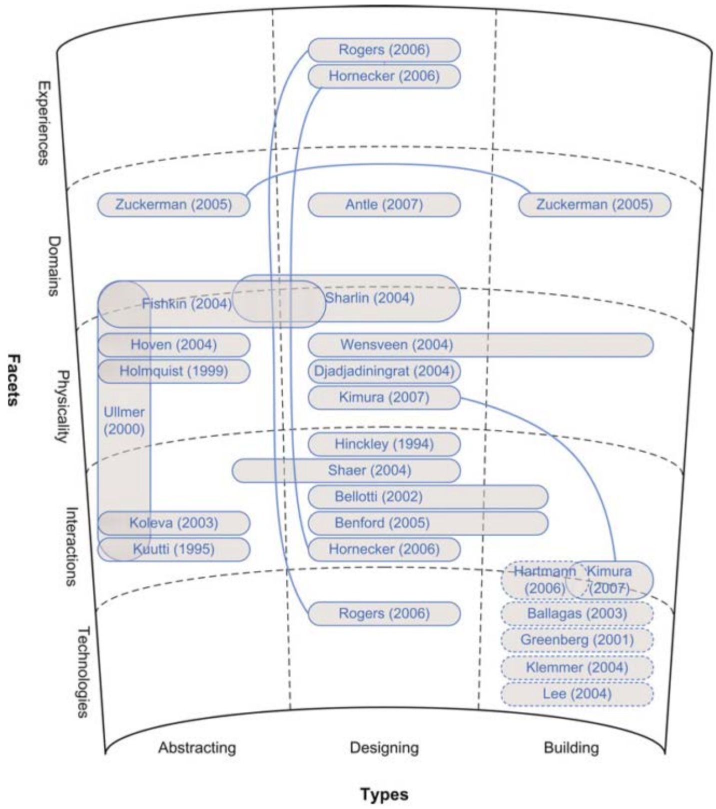

Figure 4. Classification of Tangible Interaction frameworks according to abstracting, designing and building types [137]

2.2.2.1 Abstracting Frameworks

Abstracting frameworks help researchers to better understand the properties of tangible interaction and propose models and taxonomies to frame the typical interactions that can be found in existing tangible interactive systems or that can be designed for future systems. Most frameworks in this category, according to the classification of Mazalek and Hoven [137], focus on the *facets* of *physicality* i.e., the artifact design and the affordances provided by object forms, and *interactions*, i.e., the interplay between the user and the digital systems as well as the links between physical objects and actions and digital objects and effects. In this section, first, I present the most popular interaction models for tangible user interfaces. Then, I present the existing taxonomies for classifying objects in tangible interaction.

The first model that helped understanding the main principles of tangible interaction was presented by Ullmer and Ishii in 2000 [205]. They opposed the Model Control Representation physical/digital (MCRpd) of Tangible User Interfaces (TUIs) to the Model View Control (MVC) of Graphical User Interfaces (GUIs), showing that in TUIs digital information has a physical representation, which is not present in GUIs, and that the physical representation is also the control (Figure 5). Indeed, the users can benefit of the tangible artifacts to have an immediate understanding of the state of the underlying digital information and they can also directly manipulate this information by means of the tangible artifacts. The digital information has not only a physical tangible representation but also a digital intangible representation, e.g., video projection on a tabletop. Although intangible representation could not be grasped, it is generally directly coupled to the tangible artifacts and can be also manipulated through this latter.

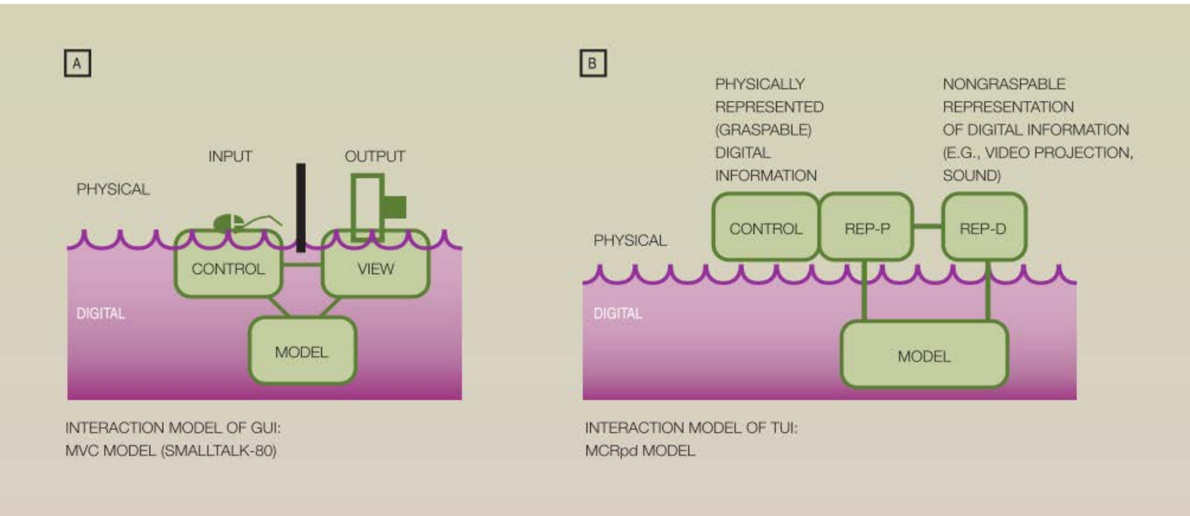

Figure 5. MCRpd model for Tangible Interaction [205]

In an updated model [97], Ishii explained that the human-computer dialogue in TUIs happens through three possible loops: in the first loop, the users' actuation has immediate feedback through the tangible representation of the information, i.e., the artifacts that they have just manipulated; in the second loop, the computer provide additional feedback by sensing the artifacts movements and by computing and updating the intangible representation; in the third loop, the computer is able to communicate with the users by actuating the tangible representation of the information, i.e., by moving the physical artifacts (and updating the intangible representation accordingly) (Figure 6). Possibly, the physical artifacts can be as small as physical pixels: in radical atoms, the actuation is done at atomic level and the users can control the form of the matter through direct manipulation but also through gestures that are not directly coupled with the matter deformation, with feedback actuated by the computer after the interpretation of the users' gestures.

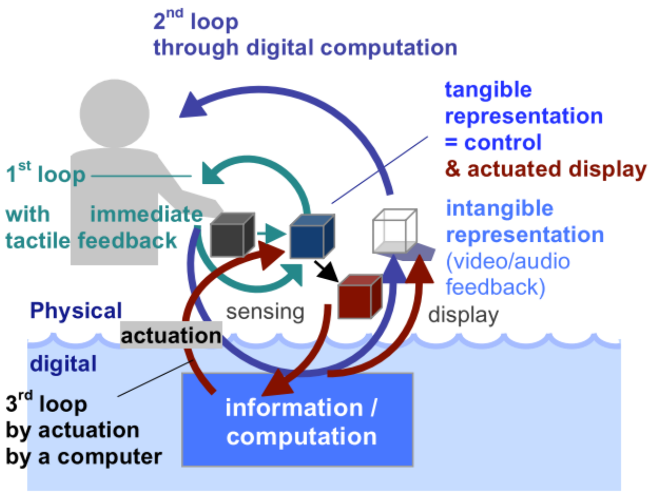

Figure 6. Interaction loops in tangible interaction [97]

Shaer et al. [187] proposed a model based on token and constraints (TAC), where tokens are physical artifacts that represent a digital information or a computational function, while constraints are all the physical elements of the interface that limit the interaction space and provide the user with affordances about how to interact with tokens. Shaer and Jakob [186] further developed the TAC paradigm describing with a formal markup language (TUIML) the interactions of Tangible User Interfaces, in particular of those that allow the manipulation of data for professional purposes. They modeled the behavior of a Tangible User Interface through a two-tier model: the *dialogue tier* is a state-transition diagram that represents all the distinct states of the application, each state differing from the others in the interactions that the user can perform. The *interaction tier* details each interaction for a particular task or function and depicts the user elementary action in the physical world and the computer responses in the digital world.

A consistent part of tangible interaction frameworks and taxonomies focused on the classification of the object types. Holmquist et al. [90] proposed a first classification of objects according to their role on the system and their physical resemblance to the digital objects. They classified objects into *containers*, which are carriers for digital information, *tokens*, which physically represent a digital object thanks to some form of resemblance, and *tools*, which allow manipulating digital information often through a computational function. Ullmer and Ishii [205] further extended this classification of objects by introducing the idea of *dynamic binding*, a property that allows the users to dynamically associate digital information to a general purpose container, and a semantic classification of objects into *iconic* and *symbolic*, whether the link with the digital object is done through physical resemblance or by an assigned symbol, respectively. Moreover, they classified systems according to the role that objects have in the system, i.e., the relationships between physical and digital: spatial, constructive, relational and associative systems. Koleva et al. [119], rather than individuating specific classes of object, classify tangible artifacts according to the degree of coherence between the physical and digital objects (Figure 7). For example, generic purpose tools that can be assigned to different functions have a low degree of coherence, while artifacts that are coupled directly to digital objects and that allow manipulating them physically have a higher degree of coherence. Koleva et al. [119] individuated several factors that can affect the degree of coherence, in particular, according to the properties of the relationship between physical and digital artifacts, e.g., whether the relationship is permanent or not (cf. the dynamic binding of Ullmer and Ishii [205]), whether the action performed on the physical objects literally corresponds to the same effect on the digital objects or a transformation is computed by the system, whether this transformation can be configured over time or not, et cetera.

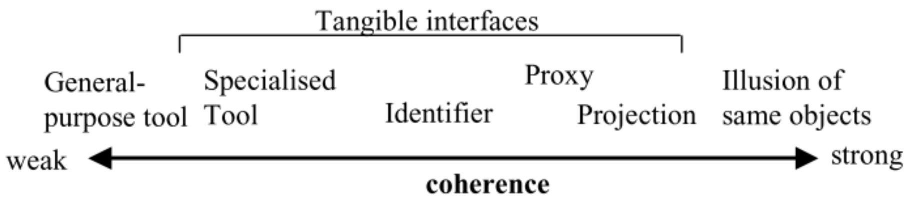

Figure 7. Koleva et al.'s categorization of Tangible Interfaces [119]

Fishkin [68] proposed a similar analysis on the objects according to two axes: the metaphors and the level of embodiment. Metaphors can be associated to objects in order to link the physical and digital world, while the level of embodiment is generally reflected by how close the system output is to the physical artifact. In particular, for the metaphor classification he built on the previous classification of Underkoffler and Ishii [207], classifying as *nouns* all the interactions in which the object or a part of the object physically resembles to the digital object and as *verbs* all the interactions in which the action physically performed on the digital object resembles to the effect generated in the digital world. It identifies then the *none* metaphor in which the object and the action performed on it have no relationship with the digital object, *noun+verb,* in which both metaphors are present, and *full*, in which the object and the interactions are so closely tied to the digital world that cannot be distinguished. Concerning the levels of embodiment, Fishkin [68] individuated examples in which the output and the state of the system are *fully* embodied in the objects and examples in which the output is presented *nearby* the object, in the surrounding *environment*, or in a *distant* display. With this classification, Fishkin [68] offers a concrete representation of the concept of embodiment in tangible user interfaces, which will be better discussed in Section 2.2.2.2. Hoven and Eggen [212] proposed a further classification of objects, extending the previous classification of Ullmer and Ishii [205] with the distinction between objects that have a personal meaning for the user and objects that have not. This classification, where meanings associated to objects are generated by the users, opens a further discussion about the design of tangible interaction that is situated in the everyday environment of the user in which particular context, such as personal experiences and culture, could affect the user interpretation of the tangible interface.

Rather than classifying the physical artifacts that take part to the interaction in TUIs, Hurtienne and Israel [96] proposed a classification of image schemas (Figure 8), i.e., the concepts derived from our previous knowledge that we unconsciously apply when we interact with our surrounding environments. They identified different levels of image schemas in a continuum, according to an increasing degree of specialization: innate image schemas are intuitive reflexes that are encoded in our DNA; sensorimotor image schemas are learned through the exploration of the external world in our early age and are generally common to people worldwide; culture image schemas, instead, derives from common practices and traditions typical of a particular region; finally, expertise image schemas, are those that human learn by practicing hobbies or by learning new skills for work; tools are particular image schemas that span across the three latter categories and are developed through the usage of specific artifacts. The spectrum of image schemas is very large (between 30 and 40 according to the authors) and are classified in 8 groups.

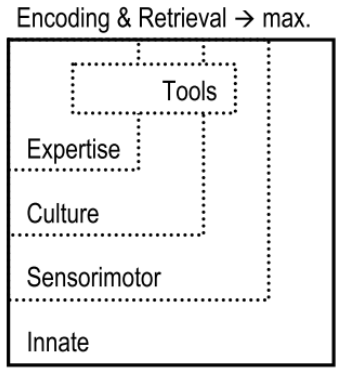

Figure 8. Image schemas classification according to Hurtienne and Israel[96]

Analyzing current abstracting frameworks on tangible interaction, most models are centered on the tangible artifacts used for the interaction, while there are few models that analyze, from a user-centered perspective, the interactions and in particular the gestures that can be performed with the tangible artifacts. While the approach proposed by Hurtienne and Israel [96] is user-centered, it focuses only on the unconscious cognitive processes that allow an intuitive use of the tangible user interface, without

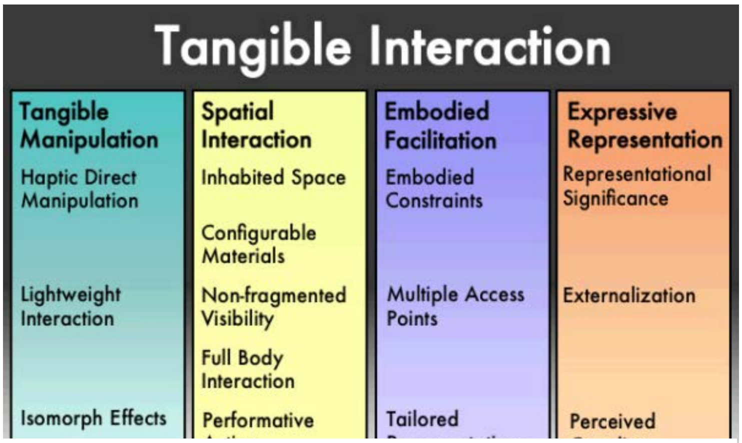

Figure 9. Hornecker and Buur's Tangible Interaction framework [95]

investigating the physicality of the gestures and the sensorimotor abilities that could be learnt to interact with tangible artifacts. Also from a semantic point of view, only Fishkin [68] considered the meaning that could be associated to gestures (with the verb metaphor).

2.2.2.2 Designing Frameworks

Designing interactions for TUIs leverage several challenges that could not be addressed by previous research on GUIs. Dourish [60], introducing the idea of embodied interaction, has been an important source of inspiration for the tangible interaction community. The concept of embodiment stems not only from the idea of embodying digital information in tangible artifacts, typical of physical computing [150] and early TUIs, but also from a phenomenological approach to the creation of meaning. Indeed, instead of thinking to the knowledge and to the meaning of interactions as existing per se, the phenomenological approach, following the theories of embodied cognition [216], suggests that the meaning is created by the humans while exploring its surrounding through their senses. From this point of view, designing embodying interactions means that the designers should provide the user with affordances that can be discovered while acting with artifacts and engaging with exchanges with other users, which involves sociological questions into the design of the interactive systems. He also advocated the necessity of a paradigm switch, with interactions that are *ready-to-hand* (following Heidegger philosopher's terminology), where the user can benefit of the system without stopping to theorize about it and to understand its meaning. In this latter case, instead, which is typical in HCI according to Dourish [60], the system becomes *present-at-hand*. Although the ready-to-hand paradigm has been criticized recently, suggesting the necessity at some point to step back from the interaction and to reflect and possibly learn new skills or knowledge [92], Dourish theory and his six principles[9] for designing embodied interaction have profoundly inspired the following trends in tangible interaction.

Several researchers provided more practical guidelines for designing tangible interaction. Bellotti et al. [25] identified five questions that the designer of TUIs (or more in general of interactions for Ubiquitous Computing) should answer: *Address*, the designer should ensure that the users can communicate easily with the system; *Attention*, the system should inform the users that it is ready to listen their commands; *Action*, the designer should establish which actions the users could perform with the system; *Alignment*, the system should acknowledge the users that it understood the users' requests; *Accident*, the users should be able to recover from an error and undo unintended commands.

[9] Note: these principles are mostly related to the aforementioned concepts and will not be detailed in this thesis for brevity purposes.

Sharlin et al. [188] evidenced three heuristics for exploiting spatiality in TUIs: physical/digital mappings should be translated in congruent spatial mappings; input and output space should be unified; the physical/digital mappings should allow the user to try different configurations, even wrong, until the right solution is found. Djajadiningrat et al. [58] propose a different perspective for the design of the interaction of tangible products: the designer should focus on direct interactions that can be performed exploiting the affordances provided by the product, rather than on symbolic interactions that rely on the previous knowledge of signs and metaphors. They also advocated a switch from a focus on visual appearance to products that are more appealing for the provided interaction. Their final suggestion is to design products that can exploits all the human skills, from cognitive to perceptual-motor skills and emotional skills. Still from a product design perspective, Wensveen et al. [224] proposed a framework for coupling action and function in tangible products through feedback and feedforward. They identified six aspects that should be coupled between user action and product reaction: time, location, direction, dynamics, modality and expression. Feedback (the information provided to the user *during* or *after* the user's action) and feedforward (the information provided to the user *before* the user's action) can be designed on three different levels: the functional, i.e., the function and services offered by the product, augmented, i.e., complementary information that exploits cognitive skills to indirectly inform about the product functions and inherent, i.e., information coming directly from the performed action through the user's perceptual-motor skills.

Among abstracting and designing frameworks presented until 2006 it is possible to notice already several different perspectives about tangible interaction. While the first frameworks and models focused on a data-centered vision and on applications for professional use, newer frameworks focused more on the user engagement started appearing. In 2006, Hornecker and Buur [95] formalized this new trend by identifying four main themes and several concepts that designers can investigate for each theme (Figure 9). The first theme is *tangible manipulation* and investigates the typical concepts of data centered systems where data can be directly manipulated through continuous interactions. With *spatial interaction,* they discussed not only the typical concepts of spatiality of Sharlin et al.[188], but also the possibility to exploit the surrounding space or full-body interactions, including performative and communicative body movements. With *embodied facilitation,* they stressed the importance of physical space and artifacts to leverage user behaviors such as collaboration, social exchanges and exploring physical skills that go beyond intuitive interaction. Finally, with *expressive representation,* they highlighted the importance of having significant physical representation of the digital data, with a perceived coupling between user's actions and system reactions and long-lasting relationships between artifacts and data that could help users' discussions. Hornecker physically implemented these suggestions for designers with a set of cards that contain questions and images concerning the different concepts and themes [93]. The card set should be used to stimulate the designer reflection during the conception of a new tangible interactive system.

A milestone framework for post-WIMP interfaces, including Tangible User Interfaces, is Reality-Based Interaction [99]. Jacob et al. [99] discussed the importance of leveraging different human skills in the interaction with real-world ubiquitous interfaces: *naïve physics*, the common understanding of the physical world; *body awareness and skills*, the knowledge of our body and of its potentialities; *environment awareness and skills*, the knowledge about the surrounding environment and the ability to interact with it; *social awareness and skills*, the knowledge about other people and the ability to communicate with them (Figure 10). Engaging these human skills is crucial in reality-based interaction, but they should be traded off with more practical considerations of the application, i.e., the *expressive power*, the *efficiency*, the *versatility*, the *ergonomics*, the *accessibility* and the *practicality*.

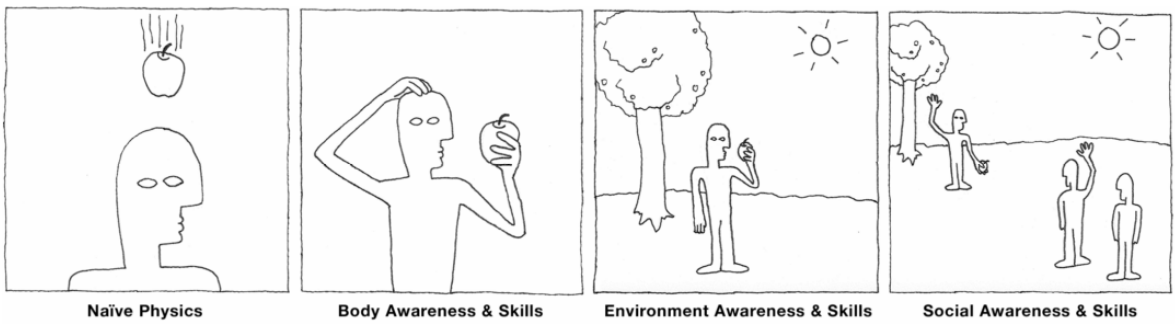

Figure 10. Reality Based Interaction Framework [99]

As already discussed at the beginning of Section 2.2, the broadening of tangible interaction themes and concepts makes a question arise: “What is then tangible interaction?”. Hoven et al. [214] tried to answer this question, without giving a sharp definition of tangible interaction, but individuating the common foundation of tangible interactive systems and the main qualities that can be found in these systems (Figure 11). The foundations are basic elements that are highlighted also on previous frameworks and models: the presence of a *physical world*, the exploitation of *human skills* and an underlying digital *computation*. The qualities of tangible interaction individuated by Hoven et al. are *direct representation and control*, i.e., the possibility to manipulate and get aware at glance of digital information; *integrated representation and control,* i.e., the action and perception spaces are merged in the physical world and integrated in the tangible artifacts; and *meaningful representation and control*, i.e., the interaction is easy to understand for the user and it is conveyed through physical affordances, metaphors or image schemas.

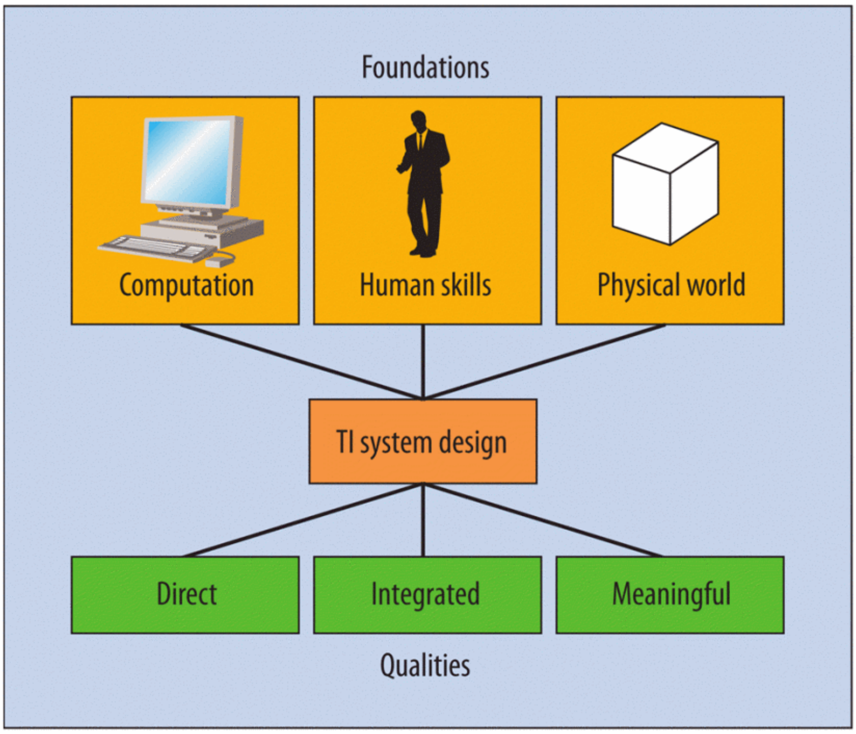

Figure 11. Foundations and qualities of Tangible Interaction [214].

2.2.2.3 Building Frameworks

This section collects the most relevant frameworks for building tangible user interfaces, consisting in software and hardware solutions that have been publicly released to facilitate the development of new tangible interactive systems. Shaer and Hornecker [185] individuated three main technological typical implementations: computer vision; RFID; and microcontrollers, sensors and actuators.

Computer vision is a popular approach to track objects on tabletop TUIs, by integrating a camera in bottom or in the top of an interactive surfaces. The TUIO protocol [105], conceived for supporting the development of tabletop interfaces and in particular of the ReacTable system [101], has become a standard for tracking blobs created by fingers and objects on interactive surfaces. Based on the TUIO protocol, the ReacTIVision Framework [104] allows to track robustly the position and orientation of several objects tagged with fiducial markers (unique visual symbols). Recently, new techniques involving 3D cameras allow tracking gestures and tangible manipulations above the surface of the tabletop. dSensingNI [118] exploits the Microsoft Kinect depth camera to detect interactions on arbitrary objects (without the need of markers) such as grasping, stacking and grouping, as well as touch interactions on the objects. The TACTIC framework [149] allows managing different inputs, e.g., a depth camera positioned over the tabletop and an infrared camera to track objects and touch gestures through the TUIO protocol. The TULIP framework [202], instead, aims at facilitating the development of tangible user interfaces by providing an abstraction layer to the tracking technologies and widgets composed by a physical handle and digital coronas directly coupled to the handle following the Ullmer and Ishii's MCRpd model [205].

RFID technologies has been adopted on tabletops to detect and track tagged objects placed on an array of RFID antennas [122]. Although, there are few popular interaction frameworks based on this technology, RFID offers several opportunities for the design of passive tangible tokens [134], which can be used also as analog dials without the need of a battery [189].

One of the most consistent technological development for supporting the development of tangible user interfaces beyond tabletop applications concerned the release of low-power small microcontrollers, sensors and actuators that could be easily integrated into objects. O'Sullivan and Igoe offer a complete guide for the integration of microcontrollers and sensors, towards a new era of physical computing [150]. To facilitate this task to people that do not have deep electronic and programming skills, several toolkits have been realized and many of them are nowadays commercially available. One of the first toolkits for rapid prototyping physical interactive systems is d.tools [86], which provides an open source hardware and software solution. d.tools sensors and actuators are equipped each with a microcontroller and communicate through an I2C bus, while the software includes a graphical editor to detect automatically the connected components and to easily manage the component behavior. Several prototyping platforms, often born in academic contexts, have been further developed and reached a commercial availability. Phidgets [80] is a very popular set of analog sensors and motors that can be controlled through a dedicated board and easily managed through dedicated software libraries. Arduino [20] and its numerous variants (e.g., the wearable platform Lilypad Arduino [32]) have become very popular in the last years thanks to the diffusion of a large community that provides examples, tutorials and direct support on the forum to various implementation of physical computing, allowing to bridge sensors, displays actuators, connectivity and many other functions on low-cost and low power hardware. The simplified programming environment allows creating simple sketches even to novice users, while it is still possible to manage more complex behaviors through C++ libraries. The framework .net gadgeteer [219], now available under the name of Micro Framework for Netduino boards [220], offers similar capabilities to Arduino but exploiting the more complete .NET language and a more powerful controller then average Arduino controllers. Littlebits [22] offers an even simplified controller that allow to program the behavior of a system through a plug and play assembly of different functional modules.

2.3 Gesture Interaction

2.3.1 Gesture Definition and Role in HCI

Since gestures are commonly used to communicate by humans and animals, there is a common understood sense in most human cultures and languages for the word "gesture". Indeed, in most dictionaries, among the various meanings, it is possible to find the common sense of a body movement (generally of the limbs) used to communicate or show something or to enhance a discourse:

- *Longman Dictionary of Contemporary English (1)*: "a movement of part of your body, especially your hands or head, to show what you mean or how you feel".[10]
- *Oxford English Dictionary (4a):* "a movement of the body or any part of it. Now only in restricted sense: a movement expressive of thought or feeling".[11]
- *Merriam-Webster (2):* "a movement usually of the body or limbs that expresses or emphasizes an idea, sentiment, or attitude".[12]
- *Wikipedia (EN):* "a gesture is a form of non-verbal communication or non-vocal communication in which visible bodily actions communicate particular messages, either in place of, or in conjunction with, speech. Gestures include movement of the hands, face, or other parts of the body".[13]

The origin of the word gesture comes from the Medieval Latin *gestura* (action) and the Latin *gestus*[14], past participle of the verb *gerere* (carry). While in the modern definition of the word there is a stronger connotation related to the movement, the origin of the word highlights also its functional role of carrying something, i.e., meaningful information.

Indeed, Kurtenbach and Hulteen [124] proposed a similar definition for HCI in 1990: "A gesture is a motion of the body that contains information." They also provided two examples to explain this definition: "Waving goodbye is a gesture. Pressing a key on a keyboard is not a gesture because the

[10] http://www.ldoceonline.com/dictionary/gesture_1
[11] http://www.oed.com/view/Entry/77985
[12] http://www.merriam-webster.com/dictionary/gesture
[13] https://en.wikipedia.org/wiki/Gesture
[14] http://www.etymonline.com/index.php?term=gesture&allowed_in_frame=0

motion of a finger on its way to hitting a key is neither observed nor significant. All that matters is which key was pressed".

Kendon [112] proposed a characterization of gestures to distinguish them from other human activities (such as practical actions and posture adjustments) based on the notion of phrases, which implies a clear movement back and forward from a rest position, with a peak that identifies the "core" of the gesture. Similarly, Schegloff [182] analyzed gestures during the speech identifying an onset and offset of the gesture, as well as the acme of the gesture, i.e., the point of maximum extension of the gesture. Similar analyses have been applied also to HCI, with a particular emphasis on gestures phrases to communicate with the computer. While developing the Charade system to allow users interacting with computer through mid-air hand gestures, Baudel and Beaudouin-Lafon [21] evidenced the need to define gestures delimited in time that can be easily segmented by the machine. In 1984, Buxton [35] already analyzed the syntax of the interaction with computer through physical gestures with the mouse, stressing that the main limiting factor was the restricted range of available gestures to interact with the computer. Defining a vocabulary of gestures easy to remember for the user and easy to distinguish for the machine from user's common movements and gesticulation is a typical challenge of gesture interaction. Still recently, Golod et al. [78] proposed a representation of the gesture phrase for microinteractions (interactions that last less than 4 seconds) individuated by peaks in the muscular tension in the activation and closure phase of the gesture phrase.

These previous works suggest that gestures can be considered as languages to communicate with computer and, therefore, they could be analyzed using the same formalisms generally used to study human language.

2.3.2 Gesture Language: Studying Gestures Through Semiotics

As already discussed in Section 1.1, there is a close relationship between gesture and languages, since they have common origins [72]. Moreover, studies of the neural systems [234] confirm that the production of symbolic gestures activates the same regions of speech production, supporting the hypothesis of the usage of common semantic constructs in the cognitive processes that generate speech and gestures. McNeill [138], analyzing the results of the study of Golding-Meadow et al. [77] with deaf children, who cannot be influenced in gesture production by the learning of a spoken language, supported also this hypothesis. Indeed, a deaf child who was unable to communicate his request to make fall a coin in the ground through a unique gesture composed by a OK shape hand and a downward movement, in order to better explain his intention to the observer, broke the aforementioned gesture into elementary components, i.e., a static OK shape of the hand to illustrate the coin, and a downward movement with a flat hand, to illustrate the fall. This observation shows that also gestures, as spoken languages, can be decomposed in atomic elements (called by McNeill morphemes). Recently, Vaucelle and Ishii [217], who can be considered the first to theorize gesture interaction with objects, stated that “Gestures scale like a language, have different contexts, different meanings and different results”. The similarity between gesture interaction and languages, which both use signs to communicate meanings, suggests that they can be analyzed under the broader discipline of semiotics. Indeed, semiotics studies the signs in their most generic form, their meanings and how they are used to communicate. Semiotics includes also the branch of linguistics, which investigates more specifically the structure and the meaning of languages.

Ferdinand de Saussure, who is considered as one of the father of linguistics, represented the sign through a signifier-signified pair [53]: for example, the word “tree” is associated to the concept of the tree, as universally known by English speakers. However, in general, the interpretation of a sign is a much more complex task than resolving a signifier/signified bijective function. Peirce [156] proposed a triadic vision of the sign, introducing the *interpretant*, which represents the user’s interpretation of the association between the *sign vehicle* (a word, a gesture, etc.) and an *object* (for example the tree). Bühler’s Organon model [33] offers an alternative triadic vision of the sign, which has three communicative functions (*expressive*, *conative* and *referential*) (Figure 12).

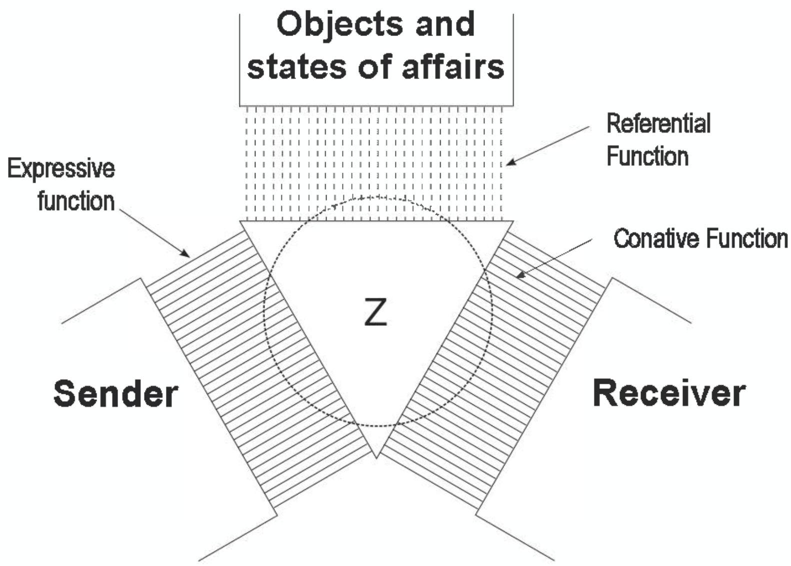

Figure 12. Bühler's Organon Model [33]

Semiotics has been applied to Human-Computer Interaction also to discuss how the interaction designer could communicate with the users through the interface. Indeed, De Souza [54] in “The

Semiotic Engineering of Human-Computer Interaction" argued that the designers should be able to convey the meanings of the designed interaction through the artifacts of the interface, such as words, graphics and interface behaviors. Sjöström and Goldkuhl [190] underlined also that users can communicate with other users through the interface, even if the interfaces was not explicitly designed for this purpose. These considerations are particularly important for the design of gesture interfaces, where it is important to convey to the users the meanings of gestures and where gestures can be used to communicate at the same time with the machine and with other users, especially in collaborative applications. The implications for tangible gesture interaction will be discussed in Section 5.4.

Generally, semiotics studies the signs through three different branches: syntactics, semantics and pragmatics. Syntactics analyzes the formal structure of the signs and how they can be combined to build sentences. Semantics focuses on the meaning of symbols as conventionally defined, while pragmatics explains how these symbols are used taking into account also the prior knowledge of the user and the context of the utterance. This classification of semiotics will be used to structure the abstracting part of the Tangible Gesture Interaction Framework, which is presented in Section 4. An analysis based on semiotics has been conducted by Kammer et al. [107] for multi-touch gestures. The authors proposed a formalization for the syntactics of multi-touch gestures on flat surfaces, with a complete grammar for the definition and composition of gestures. The authors proposed also a semantic classification of the different types of gestures, while the analysis of users' mental models, i.e., the pragmatics, was left as future work. No other comprehensive semiotic analysis has been found for gesture interaction, although several formalization of gesture syntactics and several semantic classification of gestures have been found in literature. The results of this research are presented in the following sub-sections. In particular, they present the related work in the field of gesture interaction following the same categorization in abstracting, designing and building frameworks [137], already adopted in the previous section for tangible interaction.

2.3.3 Gesture Interaction Frameworks

Since gestures can assume various form, performed through movements of the hand or other parts of the body, HCI researchers often felt the need of framing this broad design space. Hoven and Mazalek [213] considered three main types of gestures for HCI: gestures in 3-D space, gestures on surfaces and gestures with objects. Several frameworks exist for gestures in 3-D space and gesture on surfaces. Conversely, less frameworks exist for gesture with objects: this part will be analyzed in Section 2.4. As for tangible interaction, some researchers strived to provide high-level frameworks, in particular with taxonomies to help distinguishing the different gestures that can be designed to interact with computer. Others focused on technical solutions to describe gestures more formally. These formal descriptions were used for designing the interactions as well as for simplifying the recognition of gestures. In this section, I used the same categorization used in Section 2.2. Indeed, I framed gesture interaction frameworks in abstracting, designing and building frameworks.

2.3.3.1 Abstracting Frameworks

One of the first gesture taxonomies has been proposed by Kendon [113], who classified gestures that humans perform to communicate (with other humans) on a continuum (Figure 13), from gestures that accompany speech and that are less formal (gesticulation and language-like gestures and pantomimes), to gestures that require a particular formalism (emblems and sign languages).

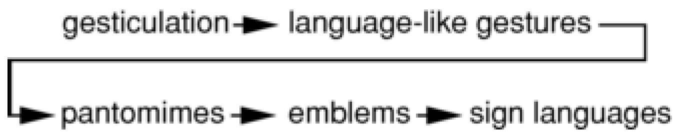

Figure 13. Kendon's representation of gesture continuum [113], from Quek [163].

Quek's studies on gestures and speech were mostly oriented to Human-Computer Interaction and in particular to computer vision techniques for gesture recognition [164]. Therefore, the hierarchical-tree gesture classification that he proposed (see Figure 14 for an adapted version) reflects the need to understand which information each gesture can carry and how this information can be interpreted through computer-vision techniques [163]. For this reason, Quek divided gestures into communicative gestures, hand movements performed with a communicative intent, generally in mid-air, and manipulative gestures, hand movements performed to manipulate objects and/or modify their properties [163]. Quek focused his analysis only on the first category, because communicative gestures were considered more appropriate for computer-vision gesture recognition techniques. Quek divided communicative gestures in two main categories: acts, which are generally closer to the intended interpretation, and symbols, which have a more abstract referential function. Acts are further divided in two classes: mimetic, which are mainly iconic, and deictic, such as pointing gestures. Symbols are also divided in two further categories: referential gestures, which have a standalone meaning, and modalizing gestures, which acquire meaning according to additional context information.

In her PhD thesis, Karam's [108] tried to offer a wider view on the gesture interaction design space. Karam classified gestures along four categories: gesture style, application domain, input (enabling technologies) and output (system responses). In particular, extending Quek's taxonomy [163], Karam identified five types of gesture styles: deictic (pointing), gesture-speech approaches (gesticulation), manipulations, semaphores, and language-based gestures (sign language). In particular, manipulations and semaphores deserve a particular attention. Indeed, according to Karam, manipulations are gestures performed to control directly either virtual or physical objects. 2D manipulations of digital objects can be performed through gestures on surfaces (such as interactive tabletops); an additional component (such as pressure) can be used to add a third dimension to the digital manipulation. In this category, she identified also tangible gestures, i.e., the hand manipulation of physical objects to control digital or physical objects. This category will be further discussed in Section 2.4. Concerning semaphores, which can be considered as a dictionary of symbols to communicate with the machine, she distinguished between static (hand or body poses) and dynamic gestures, which are performed through hand or body movements. Within the semaphore category, she considered also strokes, a particular class of movements performed on a surface with a finger, a mouse, or a pen in order to draw a particular path or mark.

A complete framework for hand gestures for HCI has been proposed by Pavlovic et al. [155]. The framework, which is mostly oriented to computer vision techniques, deals with gesture modeling, gesture analysis and gesture recognition and summarize existing applications and systems. The gesture modeling proposed by Pavlovic et al. [155] is based on existing transfer function models of speech communication. In this model, the intended gesture is transformed in a hand/arm movement, which in turn is transformed in a visual image. Therefore, the estimation of the gesture is performed through an inverse function that from the perceived visual image understand the movement and interpret the gesture. The authors proposed also an extension of Quek's taxonomy [163], differentiating between unintentional hand/arm movements, and gestures, which are then classified according to Quek's taxonomy (Figure 14). Pavlovic et al. modeled gestures also along two additional dimensions: time and space. The temporal model divides gestures in three phases: the preparation, the nucleus and the retraction. Only the nucleus (also called peak or stroke) carries the important information, while the other two phases are generally fast movements to initiate and conclude the gesture. Finally, typical spatial models for gesture recognition are based on the 3D model of the hand or arm (more recently of the whole body) or on appearance-based models that estimate gesture through parameters of the captured images (e.g., motion flow). Based on these models, Pavlovic et al. presented existing techniques to analyze gestures (i.e., extract features and estimate model parameters) and recognize gestures. Finally, they classified existing systems according to their framework.

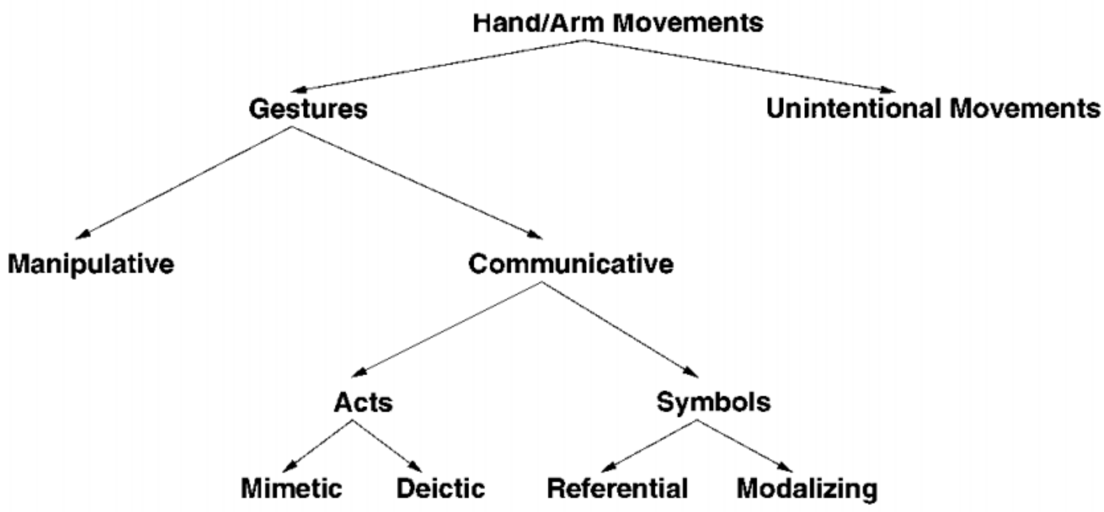

Figure 14. Taxonomy of gestures of Pavlovic et al. [155], adapted from Quek's taxonomy [163].

While the previous frameworks deal mostly with mid-air gestures in 3D space, there is a consistent body of work on gestures on 2D surfaces. Cirelli and Nakamura [48] proposed a review of multitouch gesture frameworks and gesture definition approaches. They individuated different gesture classes according to different properties: degrees of freedom, spatial, semantics, trajectory complexity, number of users, number of fingers and timing. Moreover, they identified two main approaches for gesture definition and recognition: those based on a formal definition of gestures by the designers, and those relying on user defined gestures and machine learning.

Among most popular gesture formalization, it is worth citing Proton++ [117] and GeForMT [107], two formalizations based on regular expressions to describes multitouch gesture syntax. While the original version of Proton [116] identifies only three basic gestures that can be codified as regular expressions (translation, rotation and scale), Proton++ [117] allows defining custom gesture trajectories and to take into account additional attributes, such as hand and user identification and touch shape. As discussed in Section 2.3.2, Kammer et al. [107] proposed GeForMT, a formalization of gestures for multitouch based on semiotics. The gesture syntactics is composed by different language elements, i.e., atomic gestures and additional attributes and context elements, such as, pose function (one or more fingers, or the whole hand), composition operators, focus (e.g., the touched object) and area constraints. Examples of atomic gestures are point, hold, move, line, circle, etc.

While the previous formalizations have been generated by the authors, mainly with the purpose to facilitate the development of multi-touch gesture interfaces, Wobbrock et al. [230] proposed a different approach to define a taxonomy of multitouch gesture. The authors conducted a gesture elicitation study, asking to 20 people which gestures they would like to perform for 27 different commands. The purpose of individuating a user-defined taxonomy of gestures is maximizing the guessability of gesture-command couples for the end-user. In a later study, Morris et al. [144] demonstrated that users preferred user-defined gestures to the gestures defined by HCI researchers. The derived taxonomy [230] classifies gestures according to four dimensions: form (which describes how the gesture is performed, either statically or dynamically), nature (which describes how the gesture is mapped to its referent command), binding (which describes the focus of the gesture), and flow (which indicates if feedback is discrete, i.e., shown after the gesture, or continuous, i.e., shown during the gesture).

2.3.3.2 Designing Frameworks

Bellotti et al. [25] framework is relevant not only for tangible interaction (cf. Section 2.2.2.2), but also for gesture interaction. Among the 5 “A” addressed by the framework, Address and Accident are particularly relevant in the context of gesture interaction. Indeed, it is often difficult for the users to understand where and how they can perform gestures to address the system. False positives during gesticulation and the possibility to recover from errors are also critical aspects of gesture interaction that should be carefully addressed by the interaction designer. Most of these problems were already discussed by Baudel and Beaudoin-Lafon [21], in 1993. Indeed, besides the benefit of providing natural and expressive interactions without the need of additional devices, free-hand gesture interaction raises some critical issues that should be addressed in the design phase. A typical issue of free-hand gesture interaction is associated to the fatigue that some gestures can generate in the user, the so-called gorilla-arm effect. Another typical issue that should be addressed while designing gesture interaction is how to inform the users of the gestures that they should use to interact with the system. Finally, as discussed by Pavlovic et al. [155], gestures need to be delimited in time, and should be well separated from everyday gesticulation to avoid that the system interprets unintended commands. Baudel and Beaudoin-Lafon [21] suggested that designers of gesture interaction should exploit hand tension (full-finger extension) to distinguish between unintentional gesticulation and commands for the system, which should be executed quickly by the user to avoid fatigue. The users should also receive prompt feedback from the system. At the same time, commands should be easily reversible to undo unintended actions and easy to learn for the user. The authors stated that particular application contexts, such navigating a menu or drawing software, are more suited for gesture interaction in respect to other applications domains, such those where symbolic mappings are needed.

Recently, Golod et al. [78] adapted the guidelines of Baudel and Beaudoin-Lafon [21] to microinteractions, i.e., interactions that last less than 4 seconds, without distracting the user from the main task. In particular, they characterized the gesture phrase with four phases: the activation through muscular tension; the support of system feedforward and feedback corresponding to gesture activation; the possibility to perform incremental actions within the same gesture phrase and the closure of the gesture phrase (Figure 15). Moreover, as in Charade [21], they suggested designing gestures that are self-revealing, that allow recovering from errors and that can reduce fatigue. Finally, they argue that gesture interactive systems in ubiquitous environments should be unobtrusive, reliable and sustainable (from an energetic point of view).

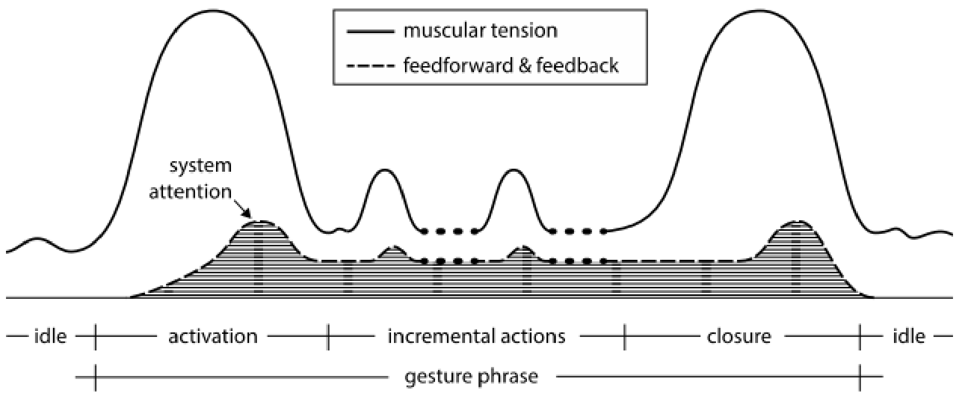

Figure 15. Gesture representation of Golod et al. [78].

In order to facilitate the learning of the gestures to the users and to reduce the cognitive load for remembering the gestures before to the execution, Carrino et al. [42] proposed an approach called "functional gestures". The aim of this approach is reducing the gesture vocabulary to interact with smart environments, associating to each gesture different behaviors according to the context of execution of the gesture. In particular, gestures are not directly associated to actions, but to functions, such as selecting, going to the next item, etc. These functions are executed on interactive elements, which can be virtual (a media-player software) or real (a lamp or a fan).

2.3.3.3 Building Frameworks

Few frameworks address the development of gesture interaction system from a broad point of view. Indeed, most frameworks focus on specific gesture recognition techniques, dealing only with free-hand gestures or with multitouch gestures. Karam [108] offered the largest vision on enabling technologies for gesture recognition, distinguishing between perceptual systems and non-perceptual systems (those that require physical contact from the user). Typical techniques for the first category are computer vision and remote sensing, such as electric field sensing. Non-perceptual systems are based on a wider range of technologies, such as mouse or stylus input, or other objects embedded with sensors, touch and pressure sensing on interactive surfaces, wearable sensors (e.g., inertial sensors), or other worn devices, such as gloves.

Few frameworks support the recognition of different gesture types and allow merging different input techniques for gesture interaction. ARAMIS, a framework developed by Carrino et al. [43, 44], supports interaction in smart environments through an opportunistic approach, allowing exploiting both environmental and wearable sensors. It provides the essential building blocks to implement gesture segmentation, gesture feature extraction and gesture recognition through machine learning techniques. The framework has been used in our lab to develop several gesture interactive systems, some of them also in the context of this thesis. Many other frameworks focus only on a specific gesture recognition technique. I already mentioned Pavlovic et al.'s framework for vision based techniques [155]. Rautaray and Agrawal [169] proposed a more recent survey of vision-based techniques for gesture recognition and typical application domains.

Kammer et al. [106] proposed a taxonomy of frameworks for building multitouch interfaces. The authors classify existing frameworks according to the underlying architecture to track fingers and manage events, according to the scope or main purpose, i.e., tracking fingers, recognizing gestures, or interacting with tangible objects on the table, and additional features, such as the possibility to add new gestures to the vocabulary or the possibility to add visual controls to the interface. Among the frameworks that support the development of multitouch gesture interfaces, I already mentioned Kammer et al.'s GeForMT [107] and Proton++ [117] frameworks, which use declarative gesture definitions to support the automatic recognition of gestures, based on multitouch events. GestIT [194] is probably the only framework that supports both multitouch and free-hand gestures. As Proton++ and GeForMT, it proposes a formalization of gestures, based on a declarative compositional language.

Many other techniques have been explored for gesture recognition. Liu et al. analyzed the use of accelerometers [132], Saponas et al. explored the use of Electromyography (EMG) sensors [179]. Combined input techniques have been also explored: Zhang et al. [236] combined EMG and accelerometers, Joselli and Clua [102] proposed a framework for touch and accelerometer based gestures on mobile devices. The list of all gesture recognition techniques can be very long and this section is not intended to be exhaustive on this topic. While this section analyzed frameworks that are focused only on gestures, the next section presents frameworks that are focused on gesture interaction with objects.

2.4 Tangible Gesture Interaction

Although gestures with objects have been adopted in a number of interactive systems, few proposed a formal analysis and characterization of these gestures, or generic guidelines to design gestures with objects. In this section, I collected most relevant frameworks for tangible gesture interaction. As for tangible interaction and gesture interaction, related work is framed in abstracting, designing and building frameworks, following the categorization of Mazalek and Hoven [137]. While in the previous section, only most relevant work for each domain have been selected, here only few works [213, 217] address directly tangible gesture interaction. Indeed, in this section, I collected also works that addressed gesture interaction with objects in a wider context, or for specific application domains or gesture types.

2.4.1 Abstracting frameworks

In 1997, Robertson [176] proposed a specific characterization of embodied actions for Computer Supported Cooperative Work (CSCW). The proposed taxonomy, generated through the observation of designers collaborating to the conception of a computer game, individuates three categories of embodied actions: actions performed in relation to physical objects, to other bodies, or to the physical workspace. Physical objects are highlighted as an important resource for the users and moving these objects is a typical action that can be carried out by the users. Robertson highlighted that other typical actions with objects are constituted by gestures that are performed to communicate with other users, in particular, to emphasize particular properties of the object.

In her PhD thesis, Karam [108] described tangible gestures as a particular class of manipulative gestures, i.e., gestures performed with physical objects to manipulate physical or digital objects. Among the examples, she reported systems where the users can manipulate a digital object through a physical artifact (e.g., Hinckley et al.'s manipulation of doll's head to visualize brain sections [88]) and systems where the manipulation of a physical object generates, through computer interpretation, the movement of another object (for example a robotic arm). She also considered as tangible gestures free movements of the body performed to control an environmental output, such as in artistic performances, where computer interprets the dancer movements (gestures) to augment their performance with a digital output (e.g., music or light projection).

The first formalization of gesture interaction with objects can be dated to 2008, when Vaucelle and Ishii defined Gesture Object Interfaces [217]. The authors used a semiotic square to oppose Gesture Object Interfaces to the classical GUIs. Indeed, in Gesture Object Interfaces, the user performs gestures while manipulating physical objects, in respect to GUIs, where neither gestures nor object manipulations are exploited to interact with the information. Gesture Object Interfaces differ also from Gesture Recognition, in which the hand of the user control directly the digital information, and from Tangible User Interfaces, in which the information is directly controlled through the manipulation of physical objects. Gesture Object Interfaces, instead, are intended to animate objects with meaningful gestures, introducing what they called "an identity reinvention". For Vaucelle and Ishii, the powerful metaphors associated to gestures can animate or give a new identity to objects. At the same time, interacting with physical objects facilitates spatial cognition, which is absent in free-hand gesture interaction.

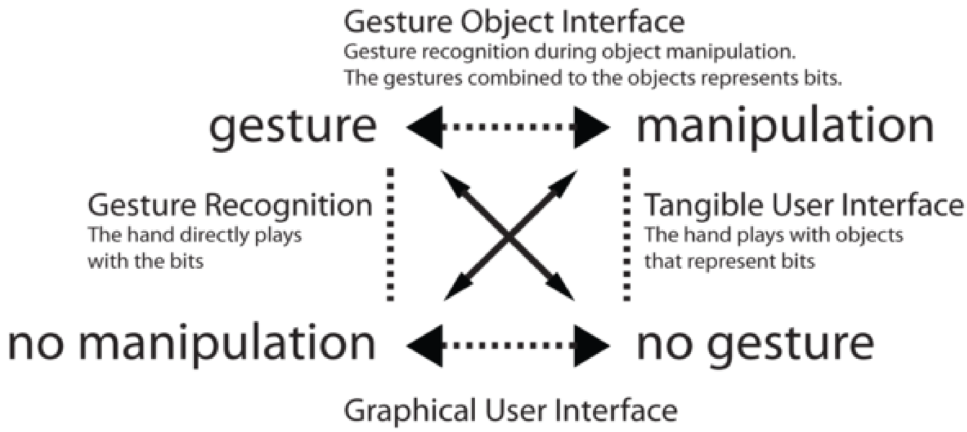

Figure 16. Gesture Object Interface semiotic square by Vaucelle and Ishii [217].

In 2011, Hoven and Mazalek [213] presented the first definition and characterization of Tangible Gesture Interaction (TGI): "the use of physical devices for facilitating, supporting, enhancing, or tracking gestures people make for digital interaction purposes. In addition, these devices meet the tangible interaction criteria". In their paper, they presented also a review of gesture interaction, distinguishing between (free-hand) gestures in 3D space, gesture on surfaces, and gestures with physical objects, specifying that all systems in the latter category can be considered as examples of tangible gesture interactive systems. In this category, they found examples of gestures with mobile devices, gestures with batons and wands, gesture with game controllers and remotes, gestures with dolls, toys and props and, finally, gestures with custom tangibles. For this latter category, Hoven and Mazalek noticed that researchers often build their own objects with which the user can interact through gestures, or they build small sensors that can be attached to existing physical objects to transform them in interactive objects.

Besides Vaucelle and Ishii's [217] and Hoven and Mazalek's [213] ground works on tangible gesture interaction, some frameworks have been proposed on specific gestures with objects. Wolf et al. [231] proposed a taxonomy of microgestures, gestures to be performed as secondary tasks while grasping another object. The proposed taxonomy has been elicited by experts according to ergonomics criteria, considering three different types of grasps: palm (a wrap grasp around the steering wheel), pad (a precision grasp for a credit card) and side (a dynamic tripod grasp for writing with a pen). The experts found 17 different gestures for the palm grasp, and 2 gestures for the pad and the side grasp. Additionally, each gesture can be performed with different fingers and some gestures can be performed with different acceleration and duration (obtaining a touch, a tab, or a press, according to these parameters).

Figure 17. Wolf et al.'s [231] taxonomy of microinteractions for palm grasp (A), pad grasp (B) and side grasp (C).

Wimmer [228] proposed a framework for grasp, which could be used either for implicit (e.g., activity recognition) or for explicit (e.g., gesture recognition) interaction. Wimmer considered grasps as static and stable hand configurations on an object, ignoring all the complex hand and finger movement that are necessary to reach this configuration. The proposed GRASP model aims at describing the meaning of a grasp through five elements: Goal, Relationship, Anatomy, Setting, Properties. Goal deals with the purpose of the grasp from a user perspective: a grasp could be performed either with a primary purpose or for a supportive purpose within another task and could be used to communicate either implicitly or explicitly with computer. Relationship addresses all the non-physical properties of the user and the object to grasp, i.e., the personal beliefs towards an object. Anatomy relates to the particular physical properties of the user's body, which can differ from person to person. Finally, setting and properties relate to how the object is positioned in the physical context and which physical properties it spots (surface texture, size, form, etc.).

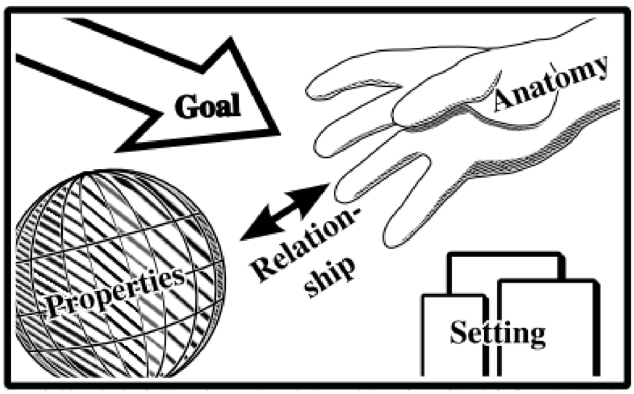

Figure 18. Wimmer's GRASP model [228].

Valdes et al. [209] proposed a taxonomy for tangible gestures with active tokens. Using Sifteo cubes, small interactive cubes with a touchscreen and inertial sensing, they conducted a gesture elicitation study to understand which gestures users would use to make queries in a large database through interactive tokens. The taxonomy classifies gestures with active tokens around three dimensions: space (of the interaction), flow (continuous or discrete) and cardinality (number of hands and tokens used for gesturing). They noticed that most interactions with active tokens were performed leaving the tokens on the surface of a table (or interactive), while about one third were performed in air. Similarly, two third of gestures were intended for discrete commands while for the remaining third the users expected continuous feedback. Finally, they noticed that the 85% percent of gestures was executed with only one hand and with only one token, while only few gestures exploited bimanual interaction or the combination of two or more active tokens.

2.4.2 Designing Frameworks

Some design guidelines for tangible gesture interaction have been found in the frameworks presented in 2.4.1. All design guidelines found in the literature review were basic and often specific for the particular type of tangible gesture.

Hoven and Mazalek [213] provided general design guidelines for tangible gesture interaction. Their first suggestion is to fit the design of tangible gesture interaction to the particular context of use, according to the particular application domain and the related target users. According to the application domain, the users could require more or less accuracy in the gesture recognition, which influences the choice of the technology for detecting and recognizing gestures. Because of the novelty of the domain, Hoven and Mazalek recommended to explore new solutions and new different contexts of use.

Wolf et al. [231] guidelines are specifically oriented to design microgestures for secondary tasks, exploiting the free attentional and motor resources of the user while performing less demanding or automated primary tasks, such as driving. The provided taxonomy helps the designer to choose gesture according to ergonomics consideration, while the choice of technology for gesture recognition should be adapted to the different types of gestures, suggesting EMG for pressing gestures and accelerometers for tapping.

Valdes et al. [209] discussed some implications for the design of gesture interaction with active tokens obtained from the gesture elicitation study. They suggested exploiting interaction beyond the surface on which tokens are placed (often a multitouch interactive table for data visualization), as well as providing continuous feedback for some of the interactions with the data through active tokens. In relation to tangible gesture design, they suggested reusing gestures across multiple commands to facilitate learnability and memorability of gestures. Moreover, they suggested designing custom active tokens to provide physical affordances, metaphors and constraints to the users. In a later work, Okerlund et al. [152] attached physical objects to the active tokens and discovered that although users had more difficulty to figure out the interaction compared to intangible representations in the token, they engaged longer in the interaction, with a better understanding of the task (biological experiments in a museum exhibition).

2.4.3 Building Frameworks

Building frameworks for tangible gesture interaction are specific for the different types of gestures. Wimmer's GRASP framework [228] provides directives for building grasp sensing systems. Three steps are required: capture, identification and interpretation. Different techniques allow capturing grasps. The user hand can be tracked with computer vision techniques to estimate the 3D joints of the hand, although often it is difficult to assess whether there is contact between the hand and the object. Data gloves with pressure sensors on the fingertips allow estimating also force contacts with the objects, but reduce the sensitivity of user's tactile perception. RFID readers attached to the user's wrist allow to detect when the user hold an object in the hand but do not allow determining the particular grasp. In alternative, the surface of the object could be instrumented with capacitive sensors or pressure sensors. In both latter cases, the capture step has the purpose to obtain the grasp signature, i.e., the digital representation of contact points between the hand and the object or the digit positions. The computer should classify the signature using either heuristics or machine learning. Finally, once grasps are identified, they should be interpreted by the system to produce some output.

Ferscha et al. [64] proposed a framework to recognize orientation gestures with artifacts. Exploiting an accelerometer that can be attached to the user's hand or to an artifact, the authors proposed a gesture library to recognize three types of gestures (32 gestures in total): orientation gestures of the user's hand, orientation gestures of a small artifact that can be held in the hand, orientation gestures with larger artifacts, which are manipulated occasionally by the users.

Klompmaker et al. [118] proposed dSensingNI, a framework for detecting interactions with object based on depth sensing (Kinect). The framework is able to detect finger touches on arbitrary surfaces (including objects); finger, hand and arm gestures; arbitrary physical objects; object-object interactions (such as grouping and stacking) and hand-object interactions (such as moving, grasping and releasing). Moreover, the framework is robust to occlusion caused by hands over the objects.

2.5 Conclusion

The analysis of existing frameworks in the domain of gesture interaction and tangible interaction allowed identifying an important body of knowledge for these two research domains, while less frameworks have been found for gesture interaction with objects. This lack is evident also from Figure 19, which resumes the previous frameworks analyzed in Chapter 2. The reason of this lack can be identified in the recent history of tangible gesture interaction as a new stand-alone research domain. Although gestures with objects have been always used to interact with the machine, it is worth analyzing their properties from a new point of view, taking advantage of the existing knowledge about tangible interaction and gesture interaction to understand better which are the benefits of blending these two well-studied interaction modalities into a new one, which still needs much exploration. On the other hand, this analysis allowed exploiting previous findings on tangible and gesture interaction to derive a framework that is grounded on existing theory in these two domains.

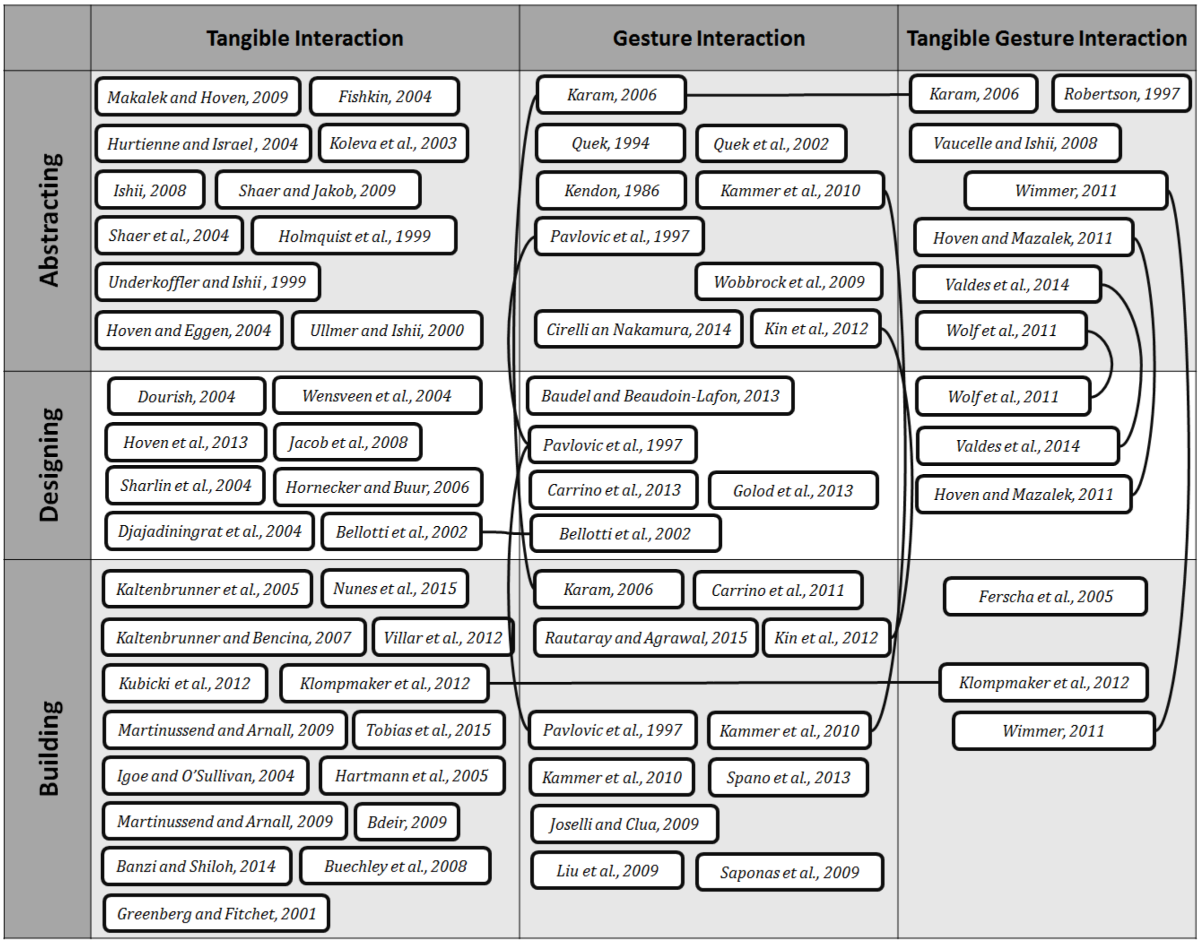

Figure 19. Classification of previous frameworks analyzed in Chapter 2.

Each domain offers particular benefits and an additional aim of the literature research was to identify which of these benefits are particularly relevant for tangible gesture interaction. A particular advantage offered by tangible interaction is the possibility to exploit the physicality of the real world, benefiting of object affordances for designing intuitive and meaningful interactions. Indeed, while the physical properties of an object can constrain and guide the type of interactions, its physical form can suggest meanings to the user, through physical resemblance, metaphors or personal memories. Indeed, most tangible interaction taxonomies focus on the functional role or on the semantic associations of the objects involved in the interaction. A key purpose of tangible interaction is also exploiting all those human skills that are generally often forgotten in traditional interfaces, such as the ability to manipulate objects and using the full body to interact with the surrounding world.

Gesture interaction literature, instead, evidenced the variety of gestures that can be performed, providing a rich vocabulary that can be used to interact with computer. Gestures are particularly flexible: the users can help designing more intuitive and easy-to-remember gestures and user-defined gestures can be used also to customize the interaction at individual level. From a technical perspective, often systems are able to recognize a large number of gestures without increasing the physical complexity of the system (adding other sensors, devices or physical objects). However, the number of gestures that users can remember is generally limited. Typical gesture taxonomies are oriented towards gesture semantics, although several frameworks try to formalize also the syntax of gestures, especially when gestures are constrained to 2D surfaces. Gesture segmentation and description in time is also a critical aspect that is often taken into account by gesture interaction frameworks.

With the diffusion of smart objects and the Internet of Things, there is an increasing interest in defining richer interactions with computer that allow exploiting the affordances of physical objects. Tangible gesture interaction, merging the benefits of tangible interaction and gesture interaction, is an interesting paradigm that could be applied to several future interactive systems. However, since Hoven and Mazalek's paper [213] is the only high-level reference for this field, there is a consistent lack of guidance for this new domain. Even if some frameworks that guide the design and development of specific gestures with objects have been individuated in Section 2.4, there is no clear link between these frameworks and the bigger umbrella of tangible gesture interaction. Moreover, even considering all those frameworks together, they are not able to cover the whole design space of tangible gesture interaction. As a result, there is a need to define a reference framework that is able to describe and support this broad design space, starting from better defining the types of tangible gestures that could be used for human-computer interaction, towards supporting the design and development of new interactive systems that exploits this heterogeneous set of signs.

From the frameworks found in literature, I selected the most relevant elements that would be used as foundation for the definition of a framework on tangible gesture interaction. Hoven et al.'s tangible interaction criteria [214] are particularly interesting to update Hoven and Mazalek's definition [213] of tangible gesture interaction (Section 4.1); previous studies on semiotics can serve as a ground for defining a communication model for tangible gesture interaction (Section 4.2); the different semantic taxonomies of gestures as well as Fishkin's taxonomy [68] of object metaphors could be unified to derive a taxonomy of tangible gesture semantics (Section 4.4.). Many guidelines for tangible interaction and gesture interaction are still valid in tangible gesture interaction (Section 5). Existing techniques for building tangible interactive systems and for recognizing gestures can serve as inspiration for building tangible gesture interactive systems (Section 6). Finally, existing frameworks on particular tangible gestures deserves the attention of a more comprehensive framework that should serve as reference to existing work as well as a tool to compare different types of gestures with objects.

3 Tangible gesture interactive systems

“ L'acqua che tocchi de' fiumi è l'ultima di quella che andò e la prima di quella che viene. Così il tempo presente.”

Leonardo da Vinci, “Scritti letterari, Pensieri, 35”

This chapter presents the state of the art for selected tangible gesture application domains, i.e., the interaction in smart homes, controlling the infotainment system of the car and the interpersonal and emotional communication. Therefore, the content of this chapter constitutes the related work for the tangible gesture interactive systems that have been developed during this thesis, which are presented in Chapter 7.

3.1 Application Domains

Several application domains are particularly suitable for Tangible Gesture Interaction. Hoven and Mazalek [213] reviewed several systems that exploit gestures with objects as interaction means with computer and individuated some typical application domains. Since gestures are strictly related to speech production, the applications where communication between users is important, such us *collaborative* and *educational* applications, as well as systems for supporting *collaborative design*, will benefit particularly of tangible gesture interaction. Hoven and Mazalek [213] found existing applications of tangible gesture interaction also for *mobile devices*, *entertainment* and *gaming*. Tangible gesture devices for gaming and entertainment (e.g., the Nintendo WiiMote and Sony PSMove) act often as remote controllers, with direct mappings between gesture and commands. Hoven and Mazalek [213] suggested that more expressive gestures could be explored, such as those used for *artistic performances* or *musical production*. During my thesis, I individuated other interesting application domains for tangible gesture interaction: the *interpersonal and emotional communication*, the *control* of devices in different contexts and the individual and collaborative *production of digital content*. Those domains have been investigated with the applications presented in Chapter 7.

Although during my thesis I reviewed several tangible gesture interactive systems from different application domains, this chapter presents only the related work that is relative to the system developed during this thesis. The choice of the application scenarios is motivated by the goal of defining a natural and economic language to interact in our everyday environments, in the context of the thesis project "Living in smart environments". Indeed, the selected application domains allowed to cover the typical environments where people spend most of their time, i.e., at home, at work, in social spaces or in the car while driving. More in detail, this chapter will focus only on work related to the following typical application domains: *remote control of entertainment systems at home*, *control of the infotainment system of the car* and *interpersonal and emotional communication*. It is worth mentioning that the state of the art for *digital content production* will not be discussed in this chapter, since the system developed for this application domain, which has been conceived for the work environment, will be discussed only briefly in this thesis (cf. the smart pen system in Section 7.6). Besides the exploration of different interaction environments, the different scenarios allowed also to cover different types of interactions: while, in the car, distraction should be minimized and the interaction should happen in the periphery of the user's attention, the emotional and interpersonal communication emphasizes the expressivity of tangible gestures and asks the users to focus on the gesture performance.

While this chapter presents only works falling in the aforementioned application domains, other examples of tangible gesture interactive systems will be presented in Section 4.3 to illustrate the different types of tangible gestures and will be classified in Section 8.2.1 according to the proposed framework. Moreover, in Section 5.4, the application domains for tangible gestures will be further discussed, presenting the common practices for the design of tangible gesture interaction.

3.2 Interacting in Smart Homes

In the context of the smart home, several researchers are investigating tangible interaction and gesture interaction as solutions for controlling the infotainment system or other household devices. Although knobs, buttons and traditional remote controllers are still the most popular solutions, gesture interaction and tangible interaction could offer richer experiences to the user, exploiting spatiality and benefiting of more human skills than the few ones required by traditional interfaces.

There are several tangible and gesture solutions for controlling the smart home. In this chapter, I will present only those who closely influenced my research, starting from the systems developed by the research groups to which I am affiliated. Accessing digital media in traditional interfaces generally requires browsing through a list of several files relying on their names or through visual previews. The Memodules project [146], developed at the HEIA-FR&UNIFR, boasts a tangible interface based on RFID-tagged objects as shortcuts for digital media, such as photos or videos. Memodules takes advantage of physical objects, e.g., souvenirs bought during the holydays, as tangible cues for the user's memories and physical links to access the related digital content, e.g., photos from the holydays where the souvenir was bought. Following the idea of Hoven and Eggen [212], the system exploits personal objects that have a special meaning for the user as graspable cues to recollect memories. Indeed, Hoven and Eggen [212] distinguished personal objects according whether they have a particular meaning for the user or not. The first category is very common in home environments and many objects can assume different meaning or recall different memories depending on the user. Moreover, personal objects can assume several forms: Nunes et al. [148] distinguished between individual or group collectibles (such as items collected to represent events or a places either personally visited or visited by the friends), worn/consumed objects (such as clothes, accessories and food that are representative of a particular place), objects to commemorate personal accomplishments (such as sport or completion trophies) and trip output objects (such as tickets, maps and pamphlets used during a trip).

Although Memodules [146] exploits the physicality of the objects and leverages the human cognitive skills that associate memories to the image of a physical object, it still constrains most of the interactions with digital media to a dedicated console. Indeed, the console integrates an RFID reader to recognize the tagged object and the users can browse the media content associated to each object through touch buttons integrated in the console. Therefore, the interaction area (cf. also Section 6.1.4) is limited to this console and the user cannot fully exploit the spatiality of their everyday environment.

Conversely, another system developed at the HEIA-FR, explored spatial interaction to control household devices with simple gestures such us pointing the device to turn it on or off [39]. The system relies on 3D cameras to track the user skeleton and to map the direction of the pointing arm to the position of the objects. Carrino et al. [42] suggest a reduced vocabulary of functional gestures to interact with household devices through a similar system: the same gesture could be used for similar commands on different objects, thus reducing to the user the time required to learn gestures. This system requires that the user registers the position of the objects in the system in order to allow the recognition of pointing gestures, limiting the flexibility over time of the interaction. Moreover, the system has a technical limitation in terms of the spatial resolution needed to distinguish objects positioned one next to the other. Therefore, while the system is particularly interesting for the interaction with bigger household devices, it is less suitable for interacting with smaller objects, such as souvenirs positioned on a shelf. In this context, wearable solution can overcome this limitation.

In literature, several wearable devices to interact with the smart home through gestures have been developed. Starner et al. [195] used a pendant to recognize free-hand gestures. The Gesture Pendant is probably the first system that used a personal worn device to recognize hand-free gestures for interacting in a smart environment. The authors stressed the advantages of portability and privacy for the user in comparison to an approach based on several cameras distributed in the environment, which often suffers from the problem of occlusion. Moreover, they showed that such system could significantly reduce the number of gestures necessary to interact in the home, exploiting the user location information. The ReachMedia project [63] has been the first attempt to benefit from the specificity of the objects for gesture interaction through a wearable device. The authors proposed a bracelet equipped with an RFID reader and an accelerometer, which allow recognizing, respectively, RFID tagged objects and gestures. In their system, when the user grabs an object, an audio menu is presented to the user, who can browse through it using gestures.

An alternative approach to interact with the smart home is performing gestures through a handheld object. Künhel et al. [123] explored 3D gestures performed with a smartphone in the hand to control different household devices. In a survey, they individuated the most common devices that users would like to control in the home: the television, lamps, the hi-fi system and blinds. Then, they conducted a gesture elicitation study to discover which gestures they would like to perform to control the aforementioned devices. The study shows that users preferred predefined gestures for popular commands where the users obtained a large agreement, while they preferred to have the possibility to define their own gesture for more uncommon commands.

Other studies explored novel designs for remote controllers, exploring tangible interfaces. Spinning in Control [115] is a remote that exploits four rotating elements to control circular menus and data in media-center interfaces. The Peppermill [218], used a similar gesture on a peppermill-shaped device to select items in a video browser; in particular, the device was powered harvesting the energy from the same user's gesture. Slurp [237] is a device that exploits the interaction metaphor of an eyedropper to pick-up digital media elements from a display or an object in the physical environment and to transfer it to another device or object in the environment, by squeezing out the media in the desired location. Slurp allows exploring spatiality, seamlessly interacting in the ubiquitous computing thanks to the tangible affordances provided by the slurp device and to interactive locations equipped with simple IR sensors.

As a result of the literature review in the context of gesture and tangible interaction in the smart home, I can summarize the following findings:

- Tangible interaction allows exploiting the spatiality of the environment and, therefore, users' spatial cognition skills
- The physical objects that can be found in the environment can be used as tangible cues for digital information
- Personal objects can be used as cues for personal memories
- The users generally appreciate the possibility to define personal gestures
- Having a small gesture vocabulary is important to reduce gesture learning times and cognitive effort for remembering gestures while interacting with the system
- People are used to remote controllers to interact with the home, but having a different controller for each device is cumbersome

The results of the investigation in this domain will be exploited during the design and development of a smart watch to recognize gestures with objects. This system is presented in Section 7.5.

3.3 Controlling Infotainment Systems in the Car

The car environment is particularly interesting for the design of innovative human-computer interfaces. Many people spend several time in the car, traveling each day from home to work, or for displacements directly related to their work. A report of the Swiss Federation shows that in average Swiss people drove 24.4 km per day in 2010 [151]; in U.K., in average motorists spend in their car 10 hours per week [201]. To alleviate this daily tedious routine, most cars integrate advanced In-Vehicle Infotainment Systems (IVIS), allowing not only the passenger but also the driver to listen to music, make and receive phone calls, navigating on a map or benefitting of even more advanced functions. The interface of the IVIS is integrated mainly in the central dashboard to allow both the driver and the front passenger to operate it. Most past IVIS interfaces were based on knobs and buttons, but, recently, large touchscreens are replacing traditional controls. While the passenger is free to operate the interface, the driver is generally occupied in the primary task, i.e., driving. Therefore, controlling the IVIS, i.e., the secondary task, can take away the driver's motor and cognitive resources from the primary task, causing distraction and potential safety risks. Indeed, to operate an interface integrated in the central dashboard, the driver has to leave one hand from the steering wheel and to move the gaze away from the road. According to Bach et al.[17], the distraction of the driver is often caused by the loss of visual perception, when the driver looks away from the road, for example to interact with the central dashboard. For this reason, having controls for the IVIS on the steering wheel allows avoiding glances away from the road, and thus potential sources of distraction. Since several years, many car-makers adopt buttons on the steering wheel to control the IVIS. However, as suggested by Riener and Wintersberger, buttons do not allow user personalization and cannot be adapted to the content of the IVIS screen [175]. In addition, buttons are generally concentrated in a small and fixed area of the steering wheel, limiting the interaction area to a small surface [59]. To overcome these limits and to offer to the driver an increased user experience, new interaction modalities are explored in research. In the last years, several researchers have explored gesture interaction in the car as an alternative and more natural interface to control the IVIS. Indeed, Riener suggested that gesture interfaces can be easy to use and can even help increasing safety by reducing the driver's visual demand [173].

Related work on in-vehicle gesture interaction has been classified according to González et al.'s paradigm "eyes on the road, hands on the wheel" [79], which is particularly interesting for increasing safety of in-vehicle interfaces. Table 1 resumes how existing in-vehicle gesture interfaces have been designed in terms of the two axes (eyes on the road, hands on the steering wheel) and, in particular, in terms of gesture types and localization of inputs and outputs. Although the paradigm "eyes on the road, hands on the steering wheel" has been proposed to increase safety, the comparison in Table 1 should not be used to compare the different systems in absolute terms, since many others design parameters of the interface could affect user motor and cognitive resources.

Previous works exploited different locations inside the vehicle to perform gestures. Many systems explored gestures performed on the steering wheel, either on its surface [79, 121], in the surrounding space [61] or on a central touchscreen [59, 160]. These gestures are performed either while firmly holding the steering wheel [59, 79, 121], or after shifting the hand position [121, 160]; touch gestures are often performed also on touchscreens integrated in the dashboard [18] and, finally, some gestures are performed in mid-air [168]. Riener and Wintersberger explored also gestures holding the gear stick [175].

Different types of feedback have also been explored in combination with gestures. When touchscreens are used as gesture input, there is quite often a visual feedback on it, although eyes-free gestures on touch-screens can be facilitated by providing also auditory feedback [124]. For mid-air gestures recognized by a Kinect, Rahman et al. exploited both auditory and tactile feedback [168]. Another common approach for "eyes on the road" interactions is providing visual feedback with a Head-Up Display (HUD)[61, 121]. HUDs are easy to be integrated into simulated environments but they are also populating many modern cars. Special techniques have been developed to avoid sight refocusing when looking at the projected information.

System or Study	Eyes on the road (Feedback)	Hands on the wheel (Gesture locations)	Gesture types	Sensing technique
EdgeWrite, González et al. [79]	++ (Head-up display)	++ (Small pads on the steering wheel)	Thumb strokes on the small pads (while grasping)	Synaptics StampPad (capacitive sensor)
Geremin, Endres et al.[61]	- Feedback behind the wheel or in the dashboard	++ Steering wheel grasp	Index mid-air strokes while grasping	Capacitive proximity sensing
Döring et al.[59]	- (feedback inside the steering wheel)	++ Center of the steering wheel with thumbs	Thumb-swipes while grasping	FTIR rear-projected touchscreen
Bach et al.[18]	+ (auditory but also visual feedback on the dashboard)	- Horizontal touchscreen on the dashboard	Index swipes / double-tap.	Touchscreen
Rahman et al.[168]	++ (haptic and auditory feedback)	-- Mid-air	Free-hand mid-air strokes	Kinect (3D camera)
SpeeT, Pfleging et al.[160]	++ (eyes-free gestures)	- Center of the steering wheel (touch screen)	Index Swipe-Scroll for direct manipulation	Touchscreen
Koyama et al.[121]	++ (Head-Up Display)	+ Steering wheel grasp and surface	Hand/thumb swipe, hand/index tap, hand drag	Infrared proximity sensors
Riener and Wintersberger [175]	-- Screen in the dashboard	-- Gearshift	Index mid-air pointing (while grasping)	Capacitive proximity sensing
Riener er al.[174] (User elicitation)	++ (eyes-free gestures)	-- Mid-air	Free-hand mid-air strokes	Kinect (3D camera)
Wolf et al. [231](Taxonomy)	++ (eyes-free gestures)	++ Steering wheel grasp	Fingers tap and swipes, hand drags, mid-air finger strokes (while grasping)	No system implemented

Table 1. Comparison of in-vehicle gesture interfaces. ++: criteria completely met, +: criteria partially met, -: criteria not met, --: criteria not met at all.

The types of gestures that can be found in existing works for interacting with the IVIS are also quite various. For this application domain, I will analyze in depth only gestures performed on the surface of the steering wheel. Gestures performed in air while holding the steering wheel, such some of those described in Wolf et al.'s taxonomy [231] or those recognized by the Geremin system [61], will not be discussed in this section. The nomenclature of surface gestures is often not consistent throughout literature. Indeed, taxonomies of touch gestures for surface computing often differentiate swipe, flick, drag and pan, according to the speed of the movement or the behavior of the GUI. For the sake of simplicity, in this thesis all gestures that imply a unidirectional movement on the surface will be considered as "swipes". I will call "strokes" the multidirectional movements and "hand-drags" the whole-hand movements around the surface of the steering wheel performed by twisting the wrist as for the "drag fingers around the wheel" gesture defined by Wolf et al.[231]. Moreover, as Wolf et al.[231], in this thesis, I distinguish between tap and press gestures whether the contact is short or long, respectively. Differently from Wolf et al.[231], I will also consider tap and press gestures performed with the whole hand. Finally, pinch and squeeze are further distinguished from the hand press, according if the hand is grasping the wheel or not. In particular, a squeeze is performed also with the palm, while a pinch apply pressure only with fingers. Among the 17 gestures elicited by experts in Wolf et al.'s taxonomy[231], 2 of them have been adopted by Koyama et al. [121], namely hand drag and index tap. The thumb swipe of Koyama et al. [8] (called "flick" in their nomenclature) was not elicited by experts in Wolf et al.[231], although a similar gesture of "dragging the thumb around the steering wheel" was present. Gestures that imply whole-hand pressure such squeeze and hand press ("push" in their nomenclature) were not considered in Wolf et al.'s taxonomy[231], although single finger pressures were included. Gestures that require leaving the hand from the steering wheel such as "hand swipe" and "hand tap" [121] were not included too, because the experts of Wolg et al.'s [231] work were asked to elicit only gestures that can be performed with a palm grasp.

In order to recognize the different type of gestures, different techniques have been used: mid-air gestures have been recognized through the Microsoft Kinect 3D camera [168, 174], finger movements while holding the steering wheel or the gear stick have been recognized through capacitive sensing [61, 175], i.e., measuring the electric field emitted by our body. Swipes and strokes performed on the center of the steering wheel or on the dashboard have been recognized through touchscreens [18, 59, 160]. Gestures on the whole surface of the steering wheel have been recognized by Koyama et al. through infrared proximity sensors, while González et al. used small touchpads (Synaptics SmartPads) to detect strokes performed with thumbs on their small surface.

Many of the systems found in literature have been designed following a technology-driven approach for the gesture design. Including end-users in the design phase of gestures can bring benefits in terms of the guessability of the adopted gestures, as demonstrated by Wobbrock et al. [230]. In the domain of in-vehicle gesture interfaces, only few user elicitations have been conducted. Pfleging et al. [160] investigated touchscreen gestures for the control of different elements of the vehicle, which were selected by voice commands. The study showed that the 78% of the gestures were 4 simple directional swipes. Recently, Riener et al. [174] conducted a user elicitation for hand gestures performed in mid-air. Interestingly, there were few consistencies among the gestures performed by the different participants, except for the most frequent locations and the hand used for the gestures (mostly the right hand).

As a result of the literature review in the context of controlling In-Vehicle Infotainment System through gestures in the car, I can summarize the following findings:

- Gesturing in the car can be an interesting approach to interact with the infotainment system since it could lower the visual demand and the cognitive resources required by other interfaces
- Gestures performed leaving both hands on the steering wheel are encouraged since they could increase safety
- While a user elicitation study explored the design space of gestures in mid-air, there's no previous user elicitation of gestures performed on the steering wheel
- Few technical solutions have been proposed to detect gestures on the steering wheel

The results of the investigation in this domain will be exploited during the design and development of four different solutions for the design and development of tangible gesture interaction on the steering wheel. These systems are presented in Section 7.2.

3.4 Interpersonal and Emotional Communication

Interpersonal communication has a key role in the human social life. As discussed in Section 1.1, humans have developed two main communication modalities: speech (verbal communication), and gestures (non-verbal communication). The evolution of modern society, with people that are able to travel easily from one side to the other of the world, has increased the need to communicate at large distances. Several technologies have been developed in the last years to ease the interpersonal communication at distance, with a particular increase of asynchronous communication. Nowadays, computer-mediated communication has an important role in our everyday life and psychologists and sociologists argue that interpersonal interaction is rapidly shifting from the physical to the virtual world [49]. Within this context, oral verbal communication has been replaced with different forms of written communication, while punctuation has partially overcome the lack of speech prosody. On the other hand, non-verbal cues, which includes gestures, postures and facial expressions, are generally more difficult to translate in digital communication means. For this reason, in computer-mediated communication it is difficult to understand and represent emotions. To overcome the lack of facial expressions and gestures in written verbal communication, emoticons have been introduced in most messaging software. Still, messaging software are able to convey emotions only through a sensory-deprived communication means, which lacks of communication modalities (touch in particular) that are crucial in face-to-face communication. Indeed, many studies showed that interpersonal touch can have several positive effects on human behavior and well-being [71]. For instance, touch is so important in romantic relationships that Montagu stated that touch and love are indivisible [141]. Several studies showed that affective touch between partners can reduce the stress [55, 81]. Touch has been demonstrated being also capable of mediating the release of oxytocin, which is the hormone that helps couples to form lasting relationship bonds [71]. In particular, a study showed that partners who exchanged many hugs in the past have higher levels of oxytocin and significantly lower blood pressure than those who reported of not having been hugged many times [129]. Moreover, Field summed up empirical research on touch and analyzed the consequences of too little touch for socio-emotional and physical well-being in childhood and adulthood [65]. Interpersonal touch plays a crucial role also in interpersonal communication: the sense of touch provides a very powerful means of eliciting and modulating human emotions. In fact, our skin contains receptors that can elicit affective responses [210]. Therefore, touch can be successfully used to share emotional aspects of communication between people.

In this section, I will review existing work about emotional communication, with a particular focus on systems that exploit gestures as interaction means, on computer-mediated touch, and on systems that encourage and augment unmediated interpersonal touch.

The "Feather, Scent and Shaker" paper paved the way for the exploration of distant communication of emotions through tangible objects [197]. In particular, the shaker system allowed sending haptic messages at distance by shaking a tangible artifact, which causes similar movements in a distant object. In this context, the LumiTouch frame [46] was the first object that enabled an explicit interaction with touch-gestures on the photo-frame surface and, at the same, time provided a visual feedback in the environment. This form of ambient display exploited the human peripheral vision in order to provide the emotional information communicated by a distant person without interfering with the user's focal attention. The benefit provided by using an interface that allows for a richer communication of emotions over a distance was demonstrated in the cubble project [120]. In this work, Kowalski et al. proposed a system to support long-distance relationships through the communication of emotional messages coded into haptic patterns and colors. The system can be used through a tangible interface, a cube filled with RGB LEDs and a vibration motor, and a mobile app in a smartphone. The two interfaces supported three kinds of messages/gestures: selecting a color and tapping one time to send a "nudge", tapping several times to send a "tap pattern" and holding the hand on the interface simultaneously with the distant pattern to initiate a warm and gentle visual flashing. The comparison between the two interfaces showed that the tangible cube enhanced the user experience and fostered a more frequent interaction leading to emotional closeness.

Designing objects for interpersonal communication at distance and in particular for the communication of emotion is a challenging task. Often, in order to promote long-lasting and reach interactions, objects are designed with life-like behavior. In this specific field, Schmitz conducted a review of the current scientific literature and presented a detailed analysis and some guidelines for the adoption of anthropomorphism and animism as paradigms for object design [184]. Life-like objects should not only have an anthropomorphic or zoomorphic form but should also present a behavior similar to a living being. Schmitz claimed that such interactive life-like objects enable building a more intuitive and more durable relationship with the user[184]. In the literature, it is possible to find some examples of zoomorphic companions for the communication of emotions. A popular example is the Nabaztag rabbit, which communicates different kinds of information with ear postures, light and sound[158]. The Hapticat is a companion that behaves like a cat [235]; a very similar concept has been proposed by Philips for their iCat, which is a robot resembling a cat and able to display emotional information performing human-like facial expressions [211].

Since social touch is considered as an essential and primary need of human life, several researchers investigated the possibility of using computers to mediate interpersonal touch over a distance. Hugging is a common means to communicate emotions and show affection to other people. Indeed, many researchers have chosen to adopt the hugging gesture for computer-mediated communication between distant people. Back in 1997, inTouch [30] is one of the first examples of mediated haptic communication. Since then, Haans and IJssenstein reviewed several research projects that aimed at communicating touch gestures through digital means [82]. Hugging is a typical gesture that researchers tried to mediate through wearable systems: in order to simulate the sensation of being hugged, Huggy Pajama [200] and "Hug over a distance" [145] use pneumatic actuators, while HugMe [45]uses an array of vibrating motor and HaptiHug [203] stretches a belt equipped with two soft hands around the chest of the user. Hugs mediated through smart clothes became popular also among the large public: in 2006, the CuteCircuit company presented the Hug Shirt, demonstrating the possibility of sending hugs over a distance with a well-integrated and stylish smart t-shirt [50].

Although computer-mediated touch could strengthen relationships between distant people, it could not substitute the benefits of physical touch generated in face-to-face encounters. Indeed, Alapack suggests that "nothing, not a fantasy, nor 'a text vanishing at the click of the mouse', can compensate for the lack of flesh-to-flesh contact in virtual communication and/or relationships"[1]. Technology brought great improvements for the interpersonal communication at distance, however, often at the expenses of interpersonal face-to-face communication. Within this context, few systems encouraged interpersonal exchanges by recognizing and augmenting interpersonal gestures. Handshaking has been investigated in several projects to share contact information. Zimmerman [238] explored first the idea of sharing contact information, in particular with a handshake, by integrating a near-field transceiver in the shoes. More recently, Wu et al. [233] implemented a similar system with a wrist-worn device, while Ketabdar et al. [114] integrated a handshaking contact information sharing system in a smart ring. While in Zimmerman's original work [238] the physical contact between people facilitates the transmission of information, in the other two works, [233] and [114], data are transmitted wirelessly through Bluetooth. At the best of our knowledge, no previous project augmented unmediated human-to-human hugs for digital communication purposes. Figure 20 resumes gestures that are commonly used for interpersonal communication and that could be augmented and encouraged with technology in order to foster social exchanges and in particular interpersonal touch. Gestures are mapped within the proxemics [84] bubbles, in order to highlight gestures that could bring a stronger emotional connotation, i.e., those closer to the intimate space.

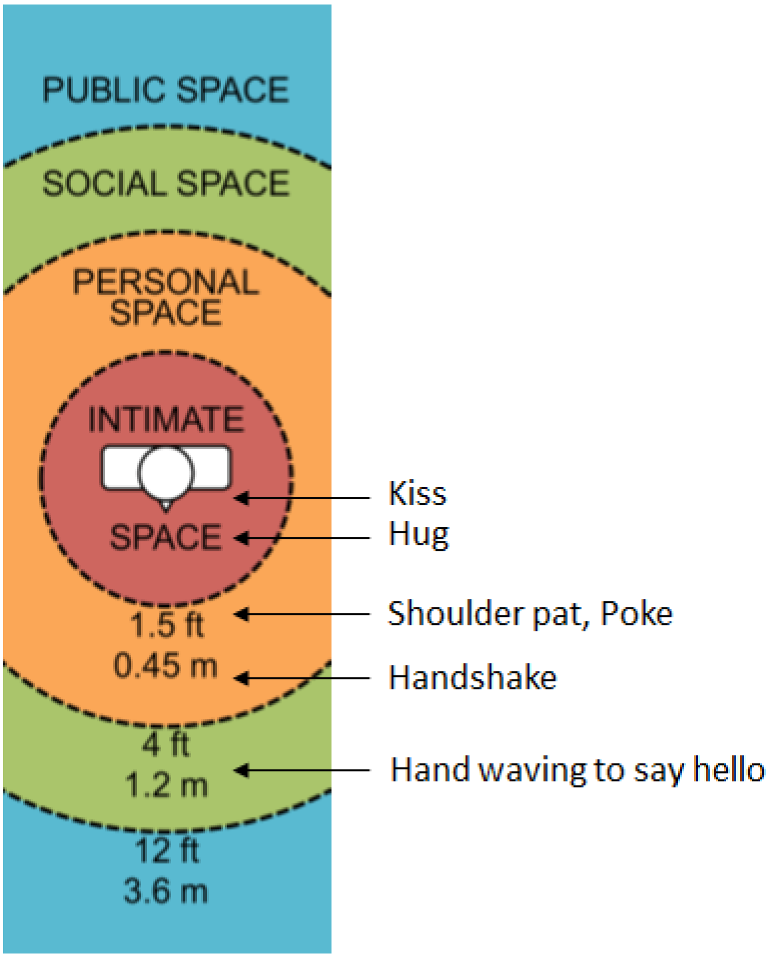

Figure 20. Gestures for interpersonal communication classified according to proxemics.

As a result of the literature review in the context of interpersonal and emotional communication through gestures, I can summarize the following findings:

- Touch has a fundamental role in the interpersonal communication of emotions
- There is an extensive body of literature in the domain of computer-mediated touch
- Life-like behaviors can enhance the emotional engagement when interacting with interactive objects
- There is a lack of systems that encourage unmediated touch gestures in the intimate space, such as hugging.

The results of the investigation in this domain will be exploited during the design and development of two different solutions, a life-like object for the interpersonal communication at distance through contact gestures and a system for encouraging interpersonal intimate touch. These systems are presented in Section 7.3 and 7.4.

3.5 Conclusion

This section presented the context and the existing work in three different application domains of tangible gestures, i.e., interacting in the smart home, controlling the infotainment systems in the car, and interpersonal and emotional communication. The research conducted in these domains obviously do not cover all the tangible gesture interactive systems that can be found in literature. Indeed, other tangible gesture types, metaphors and recognition techniques have been found in different application domains. For brevity purposes, gestures, metaphors and recognition techniques found in systems from other application domains will be described only in the respective sections. Although specific for these domains, the results found in this chapter are important to contextualize the work presented in Section 7 and in particular to understand the benefits of tangible gesture interaction in this domain.

Part II – Abstracting, Designing and Building

4 Abstracting Tangible Gesture Interaction

"Quelli che s'innamoran di pratica senza scienzia son
come 'l nocchier ch'entra in navilio senza timone o bussola,
che mai ha certezza dove si vada. "
Leonardo da Vinci, "Aforismi, novelle e profezie".

This chapter introduces the Tangible Gesture Interaction Framework and, in particular, it presents the abstracting part of the framework, which helps framing tangible gesture interactions according to the syntax and the semantics of tangible gestures. In particular, the syntax frames tangible gesture forms according to three optional components (move, hold and touch) performed in relation to one attribute and the semantics frames tangible gesture referential functions according to the coherence between the physical representation of the gesture-object pair's coherence and their meaning in the digital world.

4.1 Introduction to the Tangible Gesture Interaction Framework

Part II of this thesis presents the Tangible Gesture Interaction Framework (TGIF)[15], which aims at framing the existing knowledge and, hopefully, future works on tangible gesture interaction. The framework builds on the existing definition of tangible gesture interaction, introduced by Hoven and Mazalek [213] and already presented in Section 2.1. It is worth recalling this definition and discussing its foundations: “the use of physical devices for facilitating, supporting, enhancing, or tracking gestures people make for digital interaction purposes. In addition, these devices meet the *tangible interaction criteria*”[213]. In 2011, Hoven and Mazalek proposed Ullmer and Ishii’s framework for Tangible User Interfaces [205] as *tangible interaction criteria* for TGI. As discussed in Section 2.2.1, the research interests of the Tangible, Embedded and Embodied Interaction (TEI) community have evolved much in the last years. In this context, I argue that Hoven et al.’s [214] tangible interaction foundations (*physical world*, *human skill*s and *computing*) offer broader and more recent *tangible interaction criteria* for TGI. The three foundations can be used to characterize tangible gesture interaction and to discriminate it from other types of interactions. According to the three foundations, in TGI the users interact with objects of the *physical world*, using their cognitive and perceptual-motor *skills* for gesturing, while an underlying *computation* recognizes these gestures. This consideration helps focusing TGIF on interactions with *physical objects* that are situated in the real world rather than on interactions with virtual objects, e.g., those displayed inside a screen, and, in particular, on interactions that are consciously performed by the users to communicate with a computer using their *skills*, rather than on users’ actions performed in the real world and invisibly sensed by a computer (e.g., activity recognition). TGIF discusses in Section 4.3 the aspects related to the *physical world* and the *perceptual-motor* human *skills* required to perform gestures and it discusses the aspects related to the *cognitive skills* necessary to understand and remember the different meanings that can be associated to gestures in Section 4.4. The *computation* needed to recognize gestures is discussed in Chapter 6.

In particular, the framework discusses tangible gesture interaction according to three dimensions, i.e., abstracting, designing and building, which are derived from the three categories of tangible interaction frameworks individuated by Mazalek and Hoven [137] and builds on the previous works presented in the respective subsections of Tangible Interaction (Sections 2.2.2.1, 2.2.2.2 and 2.2.2.3), Gesture Interaction (Sections 2.3.3.1, 2.3.3.2 and 2.3.3.3) and Tangible Gesture Interaction literature (Sections 2.4.1, 2.4.2 and 2.4.3). Each dimension of TGIF is presented in a different section:

[15] The Tangible Gesture Interaction Framework has been published in [12] in 2015, redacted under the additional supervision of Elise van den Hoven. Most parts of Chapter 4, Chapter 5 and Chapter 6 are reported in their original form from this article.

Chapter 4 presents the abstracting dimension, Chapter 5 presents the designing dimension and Chapter 6 presents the building dimension. In particular, this chapter discusses and analyzes tangible gesture interaction as a language, inspired by the discipline of semiotics (Section 4.2). To this purpose, it analyzes the syntactics of tangible gestures, i.e., the forms that tangible gestures can assume and how they can be combined (Section 4.3) and the semantics of tangible gestures, i.e., the semantic constructs that can be used to associate meanings to tangible gestures (Section 4.4). This chapter, providing an overview on the different forms of tangible gestures that can be used for human-computer interaction, offers a large palette of interactions that can be source of inspiration for the designers of tangible gestures interactive systems. Guidelines on how to choose gestures from this palette are provided in the designing dimension of the framework (Chapter 5), while technical solutions for the recognition of gesture with objects are provided in the building dimension of the framework (Chapter 6). The design space of TGIF is represented in Figure A1 in Appendix A.

4.2 Communication Model of Tangible Gesture Interaction

In Section 2.3.2, I discussed how gestures can be seen as a particular language. Vaucelle and Ishii also argued that gestures with objects “scale like a language, have different contexts, different meanings and different results” [217]. Therefore, typical models of signs that are commonly used in semiotics can help understanding the role of gestures with objects in the interaction (communication) between human and computer. Bühler’s Organon model offers a rich vision of the sign, which has three communicative functions: expressive, conative and referential [33]. Inspired by the Organon model[33], the communication model of tangible gesture interaction can also be seen as a triadic function of the sign. The sign, in this case, is a tangible gesture, i.e., a gesture performed in relation to a physical object. In this model, the user performs a tangible gesture (expressive function), which is a sign expressed in the physical world with an associated meaning in the digital world (referential function); the computer receives and interprets this sign (conative function) and acknowledges the user with feedback. This communication model is depicted in Figure 21.

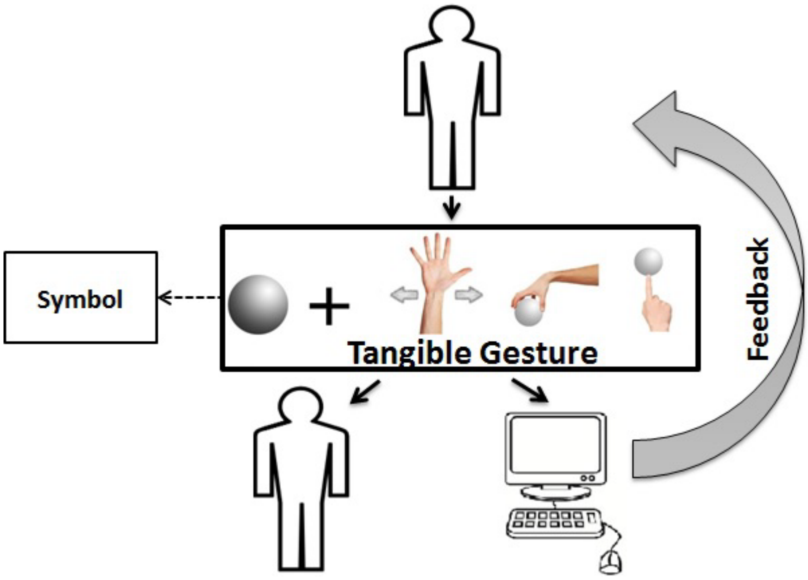

Figure 21. The communication model of tangible gesture interaction. A user performs a tangible gesture, which is a sign with an associated meaning. The computer (and possibly the other users) interprets this sign and acknowledges the user with feedback. The tangible gesture is represented as a gesture-object pair and the gesture is represented with its optional components, move, hold and touch, as discussed in Section 4.3.

Considering the model in Figure 21, TGIF focuses on two aspects: the physical phenomenon and its expressive function, i.e., the syntactics of tangible gesture, and the referential function of this sign, i.e., semantics. Syntactics and semantics are two of the three branches of semiotics, which studies signs in different domains, with some examples also in HCI [54, 107]. In semiotics, semantics focuses on the meaning of symbols as conventionally defined, while a third branch, pragmatics, explains how these symbols are used taking into account also the prior knowledge of the user and the context of the utterance. In tangible gesture interaction, semantics has a relatively short history and it is difficult to find standardized conventions, thus pragmatics generally affects most of the tangible gesture meanings. Because of the novelty of the field and the purpose of the thesis, I will discuss the two branches under the comprehensive term of semantics. It is worth noting that the communication model of Figure 21 assumes that the users and computer have a shared knowledge of the possible signs (tangible gestures) and of their meanings (effects in the system). If this common knowledge is not shared properly, the communication could be ineffective, and the system could behave differently from what the user is expecting. As suggested by De Souza for the broader field of HCI [54], the TGI designer has the important role to communicate this knowledge to the users. TGIF offers additional insights on this topic in Section 5.5.

4.3 Tangible Gesture Syntax: Move, Hold and Touch

TGIF syntax describes how tangible gestures can be physically performed by the user. To this purpose, it analyzes the physical form of tangible gestures from a user-centered perspective. Tangible gestures are pairs generated by the combination of an object of the *physical world*, and a gesture, produced by the perceptual-motor *human skills*. To characterize the form of gestures with objects, I chose to frame them according to three fundamental interaction components, move, hold and touch, as depicted in Figure 22. Although a physical object is always present in a tangible gesture, I will consider in this section the interaction with a generic object, without considering its peculiarities (represented in Figure 22 as a sphere). Obviously, objects of the physical world can have any size and shape, affecting the way the user would like to perform gestures with them according to many factors. Wimmer [228] discussed these factors for grasps, i.e., the actions (or gestures) performed to take an object in the hand. These factors will be discussed in Section 5.5 for TGI. Moreover, this section discusses gestures performed with one hand as depicted in Figure 23. Section 4.3.8 discusses how these gestures can be extended to two hands or to other parts of the body.

The choice of move, hold and touch components is well rooted in tangible interaction and gesture interaction history. Djajadiningrat et al. have presented a perceptual-motor centered approach to tangible interaction based on movements of the body as well as movements of product components [57]. Similarly, Matthews stressed the potential of using movement as rich interaction modality [136]. Price and Rogers identified three types of physicality: physical movement, interaction with physical tools and combining artifacts [162].

Fishkin et al. considered hold as an important interaction modality in their seminal work "Squeeze me! Hold me! Tilt me!" [66]. Wimmer investigated the way we hold objects as an interaction modality in [228] and [229].

Wobbrock et al. [230] and Kammer et al. [107] explored touch gestures to interact with surfaces. Touch interaction is a current trend even in commercial products, with more and more devices integrating touch gestures of different nature.

Even if holding an object implies also touching it, in TGIF I consider hold and touch as two different components. Indeed, hold and touch involve also different human haptic receptors: touching an object activates tactile receptors in the skin, while holding an object is mostly related to receptors in joints, muscle and tendons [83]. The move component, instead, relies more on vision and proprioception.

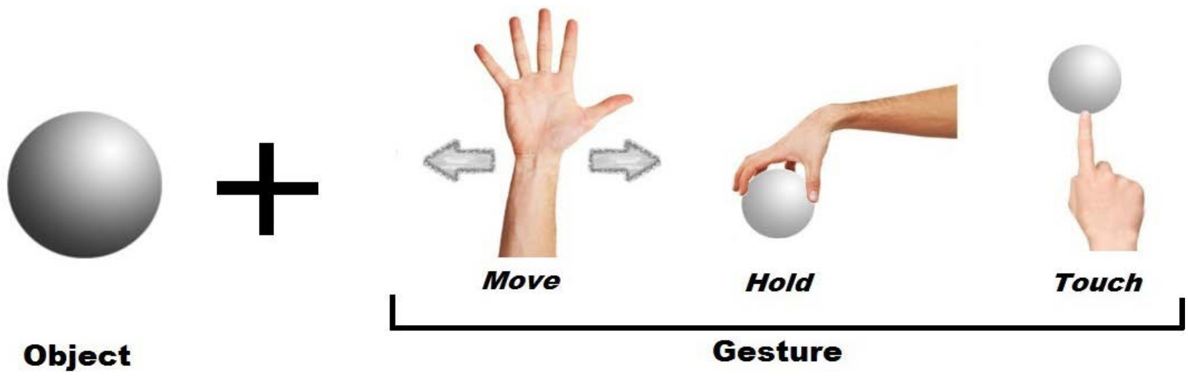

Figure 22. The Tangible Gesture Interaction Framework (TGIF) syntax of tangible gestures: a gesture based on optional move, hold and touch components, related to one object.

The purpose of decomposing gestures with objects according to the move, hold and touch components is highlighting which are the most typical types of gestures rather than obtaining a universal classification of gestures. Indeed, although the three gesture components can be combined to obtain a variety of different tangible gestures, many gestures cannot be described as a simple combination of the move, hold and touch gesture components. For the sack of completeness, I introduce attributes to characterize further gestures with objects.

4.3.1 Gesture Attributes

Move, hold and touch can be combined to obtain a rich variety of tangible gestures, as depicted in Figure 23. In the next sections, I will deepen the analysis of these various combinations with examples from literature. To enrich further the vocabulary of gestures, the designer of a TGI system can also take into account different properties that can be associated to the basic components. *Time* is a property than can be considered for move, hold and touch: how long an object is moved, touched or held can be mapped to different behaviors in the application. Similarly, the *amount of movement,* or *the amount of contact points* of a touch gesture can be used as an additional degree of freedom. The *amount of pressure* (i.e., force) can be considered as an additional property of hold and touch. Move is probably the component that offers more degrees of freedom: the designer can also consider *speed*, *direction* and *angle* as additional attributes.

4.3.2 Gesture Delimitation in Time

As in languages, where words are separated by spaces in the text or pauses in the speech, also gesture are delimited in time. Gesture segmentation is a common issue in Human-Computer Interaction, acknowledged since Charade [21], one of the first gesture interactive systems. Indeed, Human movements and objects manipulations are typically continuous and need to be separated. Golod et al. [78] delimited a "gesture phrase" with an activation and closure, both characterized by a muscular tension. A gesture phrase can be composed by one single gesture or several microinteractions to operate incremental actions.

In TGIF, tangible gestures can be often distinguished from gesticulation in an easier manner in respect to free-hand gestures, since the object provide a specific context to the execution of the gesture (cf. Section 6.3.1). Moreover, a tangible gesture is generally delimited in time by changes in a component or in an attribute of a component. For example, a static posture gesture can be delimited by movements before and after the gesture; conversely, a hold gesture can be delimited in time by the two actions of taking the object in the hand (from a no-hold state) and releasing the object (back to a no-hold state). When no clear delimitation could be individuated in a gesture, external triggers should be used to segment tangible gestures, for example speech or other gestures (cf. Section 6.3.1).

4.3.3 Single Component Gestures: Move, Hold or Touch

The simplest tangible gestures consist of only one basic component. Those are generally deictic gestures: often people touch an object or hold it in their hands to specify which object they are referring to. In Rosebud [75], a typical example of a *hold* gesture is shown: children hold toys in front of the machine in order to access stories associated to that toy. More degrees of freedom and expressivity can be added by considering additional properties of the components. One can *touch* an object in different manners: with one or more fingers, or with the whole palm. Pasquero et al. [153], for example, distinguished between touching the wristwatch face with two fingers or with the whole palm. In addition, one can hold an object applying forces, i.e., pressure, like for the squeezing gesture of tangible video bubbles [178]. Although losing the important haptic feedback given from the contact with the object, one can also perform gestures in proximity of or in relation to an object. In the ReachMedia application [63], users have first to hold an object to select its related media content and then they perform free-hand *move* gestures to browse the content. Although move, hold and touch can occur singularly in TGI, richer and more expressive gestures can be obtained by combining two or even three components. As in languages, tangible gestures components can be considered as the basic phonemes to construct complex words.

An example of the possible combinations is presented in Figure 23.

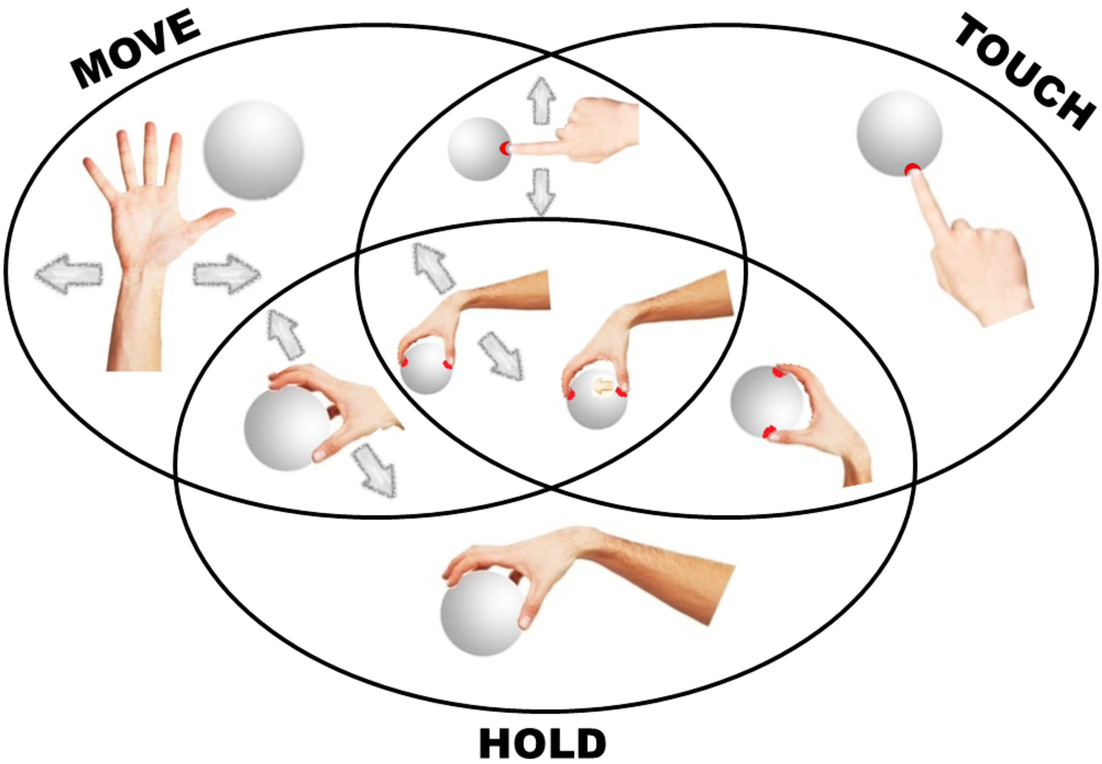

Figure 23. Taxonomy of move, hold and touch combinations.

4.3.4 Combined Gestures: Hold+Touch

As depicted in Figure 23, combining hold and touch, a grasp is obtained. A grasp is defined as the way the user holds an object in the hand, thus, differently from hold gestures, the points of the object that the user is touching are relevant to distinguish the gesture. Grip is often used in the literature as a synonym of grasp [192]. It is worth noting that grasps are static gestures and can be compared to postures in free-hand gestures. Wimmer's GRASP model [228] offers a reference frame for the design space of hold+touch gestures, with helpful guidelines also for building systems that are able to recognize different grasps. The "Human Grasping Database" is another powerful tool to understand all the physical forms that hold+touch gestures can assume [62]. Different grasps can be applied to the same object, for example, as gesture for changing the object function, like in the Microsoft multitouch pen [192] and Talyor and Bove's Graspables [199]. Grasps, therefore, are often used for mode switching [192] in TGI applications and associated with the "verb" metaphor of Fishkin [68]: one grasps the object as an "X" in order to make it behave as an "X" (cf. Section 4.4).

4.3.5 Combined Gestures: Hold+move

This class of tangible gestures includes all dynamic gestures made by moving an object while holding it. The held object has a very important role because different objects can give different meanings to the same movement. Conversely, different movements can give different meanings to the same multifunctional object: for example, the Nintendo Wii Remote and the Sony PS Move controllers are not only physically animated by gestures, but the user can, with proper gestures in a given context, "reinvent their identity" [217]. Another example of hold+move gestures is provided by Vaucelle and Ishii, who adopted in their system gestures performed with a doll to record and play movies [217]. Gestures with a physical object in the hand were first categorized by Ferscha *et al.* [64]. The Smart Gesture Sticker project, instead, showed the possibility to gesturally interact with everyday objects by just attaching a wireless accelerometer to them [15]. Pen rolling [27] is a particular hold+move gesture: in this case, the movement is associated to the held object but not to the hand. These peculiar gestures where the movement is associated to the object are further discussed in Section 9.2.

4.3.6 Combined Gestures: Touch+Move

This category includes all dynamic gestures that the users perform by touching an object, *i.e.*, touching the object and then moving the finger or the whole hand on its surface. The user can interact with a standalone object in the environment or with a body-worn object. It is worth noting that generally the user does not need to move the object, while his/her fingers or the whole hand swipe its surface to perform gestures. Wobbrock *et al.* presented an extensive taxonomy of touch gestures [230]. Although the study was focused on tabletop surfaces, thus on touch gestures that are generally performed in a GUI, it offers interesting insights for designing touch + move gestures. An example of dynamic touch gesture on a wristwatch touchscreen can be found in [14]. The user can swipe around the bezel to select the items of a calendar. Although the prototype implied the presence of a GUI, the gesture follows the typical paradigm of round watches. In fact, the item to be selected in the calendar was disposed in a circle near the bezel, like the hours in a watch. Instead of interacting on the bezel, Perrault *et al.* proposed touch gestures on the wristband of the watch [157]. In TZee, touch gestures on a truncated pyramid serve as various commands for manipulating 3D digital objects [226].

4.3.7 Combined Gestures: Move+Hold+Touch

Tangible gestures that belong to this class are the most complex. These gestures are generally extensions of the previous classes and are obtained by combining simultaneously the three interaction components, move, touch and hold. Often, they are just a composition of two simpler combinations, i.e., a grasp (hold+touch) and a hold+move gesture or a grasp and a touch+move gesture. For example, the users can grasp an object in a particular way, in order to activate a special mode of functioning, and then they can move it, performing a gesture in the air to activate a specific command. Similarly, a grasp could be followed by a touch+move gesture on the object held in a particular way. In this case, the object is static in the hand, while fingers move to perform touch gestures on the surface. In [192], it is possible to find an example of this category: Song et al., combined a cylindrical grasp on the MTPen with a swipe on the surface of the pen to define a command for turning the pages of a digital book. Wolf extensively analyzed these gestures in [231], proposing a large taxonomy of the best microinteractions to be associated to three different grasps. The purpose of Wolf's microgestures is the interaction for secondary tasks, while grasping the object is often associated to the primary task, for example, steering the wheel to drive the car. An implementation of microinteractions on the steering wheel can be found in [121].

Furthermore, in the hold+touch+move category, it is possible to find gestures that do not imply a movement of the forearm but only of the object. This could be achieved by manipulating the object with fingers, e.g., for bending an object [125]. These latter gestures often alter physical properties of the object, either in a permanent or temporary way. A particular preemptive grasp generally is still needed, but, in these cases, its only purpose is allowing the subsequent movement of the object. These particular cases will be also discussed in Section 9.2.

4.3.8 Extension to More than one Object and to Full-Body Interaction

All the gestures described above involve only one object, which can be either handheld or a standalone object that can be touched. Gestures with more than one object can also be considered for TGI. However, when two objects are combined in the hands, the relationship between them introduces many degrees of freedom to the interaction designer. Most of the implications that arise when interacting with more objects are explained by the tangible interaction principles of Ullmer and Ishii's [205], according to the spatial, constructive and relational categories. Therefore, the framework proposed in this paper aims to classify tangible gestures performed with only one object. As suggested by Hoven and Mazalek [213], two-handed interaction and full body interaction with objects should be also considered in TGI. Gestures performed with two hands on one object can be easily classified using TGIF by analyzing together the gestures of the two hands if the two gestures are symmetric, or separately if they are different. For example, the squeezing gesture on the tangible video bubbles [178] is performed with a symmetric action of the two hands. Conversely, the aforementioned combination of grasp and touch+move gesture for the turning page command on the MTPen [192] could be performed with two hands: one hand to grasp the pen and the other to perform the swipe gesture to flip pages. Other parts of the body could be also used. For example, in particular scenarios, the user could touch, hold or move an object with the mouth or with the lower limbs: the framework can be applied in the same manner depicted for the interaction with the hand.

4.4 Tangible Gesture Semantics: Meanings of Objects and Gestures

Several semantic classifications exist for objects [68, 205, 224] and gestures [68, 107, 108], but few consider the semantic constructs that can be associated to gesture-object pairs. Semantics is an important aspect to be considered in the design of tangible gestures. This facet of TGIF deals with a particular foundation of tangible interaction: the cognitive human skills needed to understand and remember gesture meanings. Referring to Figure 21, a tangible gesture is a sign performed in the physical world that generally is associated to an action in the digital system. The symbol associated to the sign is generally represented as the *model* of the digital world in Ullmer and Ishii's MCRpd interaction model [205] and as the *reference* in the Bühler's Organon model [33]. In order to have an effective communication, the knowledge of this reference should be unambiguous for both addresser (who performs the sign, e.g., the user in the top of Figure 21) and addressee (who interprets the sign, e.g., the computer and possibly the other users in Figure 21). Two approaches are possible to share this knowledge: the system designer explicitly shares the vocabulary of gestures and their relative meanings or tries to convey implicitly this information by embodying the interactions in the gesture-object pairs. In this latter case, the objects should provide enough affordances to make the users guess the possible gestures to perform with them [54, 60]. Section 5.5 offers more insights about how to convey the possible gestures through object form. Even when the tangible gesture vocabulary is communicated explicitly, having strong associations between the physical world (gesture-object pairs) and the digital world (digital objects and actions) could help decreasing the learning time and facilitate remembering the tangible gesture vocabulary. However, metaphors can break at some point and they are not always the best solution to make the user understand how the system actually works [92]. Therefore, the designer should always consider the full range of semantic constructs and choose the most proper physical-digital associations according to the application requirements.

In TGI, objects and gestures can have either a weak or a strong association with the referent. The object could be a multipurpose tool or an iconic object that totally resembles to the referred object. Similarly, the gesture could be defined arbitrarily (semaphore) or a movement that we typically perform during our everyday life (metaphor). Separating weak and strong references in only two distinct classes is generally not possible: following the approach of Koleva *et al.* [119], I will rather represent the possible semantic combinations of the object-gesture pairs in a two-dimensional continuum. In Figure 24, I mapped most tangible gestures semantic constructs that I found in literature according to the degree of coherence (or resemblance) between the physical instance of the object-gesture pair and the referenced instance in the digital world.

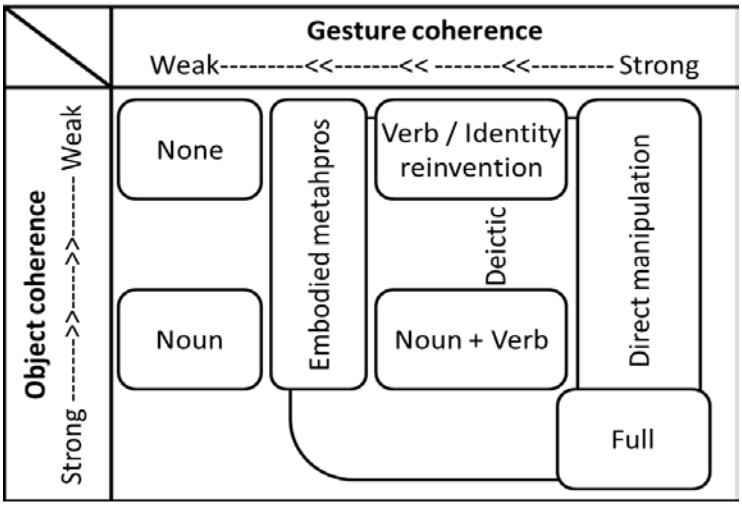

Figure 24. Map of tangible gesture semantic constructs according to the coherence between the physical representation and the referenced representation in the digital world.

To clarify each construct, I will consider the example of a tangible gesture that is associated to a function for turning on a light. Fishkin's metaphors [68] can be easily identified in the two-dimensional representation of semantic constructs presented in Figure 5. The *none* metaphor occurs when both object and gesture have a different representation than in the digital world, *i.e.*, they are both symbolic, or both arbitrarily defined by the designer. Shaking a pen (in this context a multipurpose tool) to turn on the light is an example of low coherence for both gesture and object. The *verb* metaphor has a coherent representation only for the gesture. The object can be a multipurpose tool or an object whose identity is transformed by the gesture. With the *identity reinvention*, the real object is semantically transformed in the referenced object and behaves as this latter [217]. As example of identity reinvention gesture, one can grasp a pen as a torch in order to make light into a room (cf. Figure 65 in Section 7.6).

Conversely, *noun* is a metaphor associated to objects with a strong coherence with the reference in the digital world but where the gesture has low coherence with the expected action in the system, *i.e.*, it is a *semaphore* [108, 165]. For example, the user can swipe over a lamp in order to turn on the light. For gesture-object pairs that both have a coherent reference with the digital world, Fishkin distinguishes between *noun+verb* metaphors and *full* metaphors. In the case of full metaphors, there is a complete overlap between the physical system and the digital system. An example of the full metaphor is turning (screwing) the bulb of the lamp for turning on or off the light. With *direct manipulation,* the user has no longer to translate a metaphor, nor to reason about similarities: he can directly modify the state of the system by altering some parameters. In this case, the communicative intent of tangible gestures is partially lost, however, the interaction is still explicit worth considering direct manipulation in the two dimensional continuum of Figure 24. Deictic gestures are particular gestures that are used to identify or convey the attention on a particular object. Deictic gestures generally do not have a particular meaning associated, although their forms (like pointing or taking an object in the hand) are common in our everyday gestural communication. Finally, *embodied metaphors* are particular metaphors that are based on simple concepts (embodied or image schemata [96]) derived from our everyday experience [19] and they can be used to design intuitive gestures that can be associated to objects. In direct manipulation, deictic gestures and embodied metaphors, the object could be a *container* with low coherence or could have the same representation of the digital world. For this reason, in Figure 24, they span across the whole object continuum.

It is worth noting that the reference associated to an object-gesture pair can change over time and space even in the same application, according to context information. Ullmer and Ishii [205] called this property *dynamic binding*. Moreover, the interpretation of metaphors could deeply vary among users, according to their personal backgrounds. As an example, Hoven and Eggen showed how personal objects can assume particular meaning and can serve as a link to memories [212].

4.5 Conclusion

This section introduced the Tangible Gesture Interaction Framework (TGIF) and presented the abstracting dimension of the framework. The abstracting dimension is the core contribution of the frameworks since it aims at classifying tangible gestures according to their syntax, i.e., how they are physically performed by the user, and their semantics, i.e., which semantic constructs can be exploited to associate meanings to tangible gestures. The TGIF syntax proposes a taxonomy based on the move, hold and touch components as atomic constituents of tangible gestures, which can be combined and enriched further with attributes in order to create richer and more expressive gestures. The semantics constructs to associate meanings to tangible gestures are also presented as a taxonomy in a 2D continuum. Each taxonomy is presented as a large palette of all possible design choices, without direct evidence on how it should be used to design tangible gesture interactive systems. These insights will be provided in the next Chapter. The validity of the proposed taxonomies will be discussed in Section 8.2.

5 Designing Tangible Gesture Interactive Systems

"Ogni nostra cognizione, principia dai sentimenti."
Leonardo da Vinci, "Scritti letterari, Pensieri, 31"

This chapter presents the designing dimension of the Tangible Gesture Interaction Framework. In particular, it proposes guidelines for designing tangible gesture interactive systems and different strategies for designing tangible gestures. Guidelines includes best practices for popular application domains, insights for exploiting object affordances for the design of tangible gestures and an overview of different feedback modalities. This section discusses also how to evaluate tangible gesture interaction qualities and tangible gesture interactive systems.

5.1 Overall Design Process of TGI systems

As for most interactive systems, developing a tangible gesture interactive (TGI) system requires an iterative process that encompasses both interaction and system design. While several frameworks for designing tangible interactive systems and gesture interactive systems have been already provided in literature (as discussed in Sections 2.2.2.2 and 2.3.3.2), design guidelines for TGI systems have been provided only for particular application domains or gesture types (see Section 2.4.2). In this section, I will discuss how the two-fold nature of tangible gestures (as gesture-object pairs) can be taken into account during the design process of TGI system. The general steps of this design process, which are common to most HCI disciplines, are shown in Figure 25. Following the schema presented in Figure 25, after investigating the user requirements and the aim of the application, designers should focus on the design of proper interactions that will meet the elicited requirements. In particular, after defining the functions of the systems, they should understand which types of gesture-object pairs are suitable for interacting with the desired system and which kind of output modalities can be used to provide feedback to the user. Designers should iterate several times through the choice of gesture-object pairs by testing with users the designed interactions. While the first iteration could be done with a mockup, further steps should involve working prototypes that recognize user gestures. At this point, building a working system becomes an important step to validate the design choices. Several technologies exist for recognizing the different types of tangible gestures; however, sometimes the technologies are not ready for recognizing the desired gestures with enough accuracy or power/space efficiency. Depending on the availability of technology and resources, the designed tangible gestures need to meet the implementation requirements and need to be refined or modified during the following iterations that bring to the final working system.

Within this iterative design process, the two TGI taxonomies presented in Section 4.3 and 4.4 can be adopted as reference framework for reasoning about all the possible gesture-object pairs from which designers can pick the most appropriate gesture-object pairs and about the semantic constructs that can be used to associate meaning, e.g., system functions, to gesture-object pairs. This will influence the two steps of designing the interaction and building the interactive system. The TGIF framework will also contribute to the interaction design step through the TGI guidelines presented in this chapter. Conversely, the building step is supported by Chapter 6. In particular, the choice of the technological approach to build the system is discussed more in depth in Section 6.1.4.

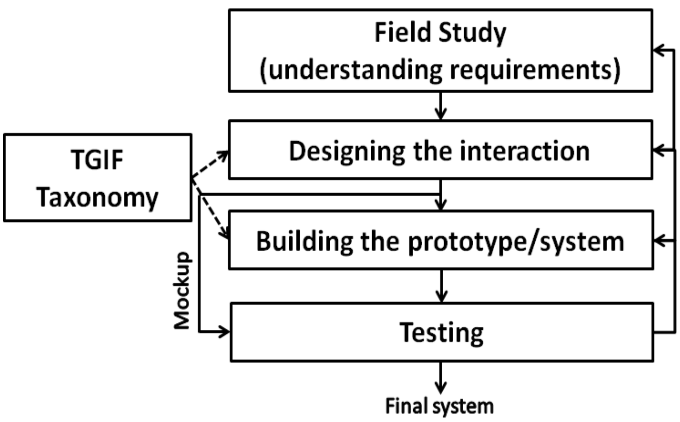

Figure 25. Design process of a TGI system.

5.2 Design methods for the definition of tangible gestures

As stressed in Section 4.2, a peculiarity of tangible gesture interaction is the adoption of gesture-object pairs as signs for the communication with computer. In tangible interaction, instead, the communicative aspects of gestures performed with objects is often neglected, while in gesture interaction, designers focused more on hand or body movements, without considering the possibility to perform these gestures in relation to an object.

In TGI, both gestures and objects have an important role and the respective design choices will influence considerably the final TGI system. For example, gestures that include the hold and touch components can increase the emotionality of the interface, while the move component can increase the expressivity and the variety of the gesture set. At the same time, the objects used for tangible gestures have an important role in the definition of the gesture components: for example, bulky objects cannot be held in the hand and in this case touch + move gestures are more suitable; deformable objects, like pillows, offer affordances for applying pressure, either while touching or holding them. Moreover, object affordances can facilitate the usage of metaphors for associating particular system behaviors to the gesture-object pairs. Since the design space of tangible gesture interaction is very large, the aforementioned considerations relative to gesture and object choices are discussed more in depth in the following subsections, which provide general guidance through examples from literature. In particular, Section 5.4 provides common practices for the choice of gesture components in relation to particular application domains, and Section 5.5 discusses the importance of object affordances in tangible gesture design.

Although the examples found in literature can help designing tangible gesture interaction, exploring new gesture-object pairs and metaphors is also fundamental to improve the interaction experience. To this purpose, two main design approaches can be followed to generate a set of tangible gestures for a TGI system: the *expert-driven* approach and the *user-driven* approach.

The *expert-driven* approach relies on the previous experience in the field of the TGI designer, on the knowledge collected during field investigations and on the general knowledge provided by TGIF. A new vocabulary of tangible gesture interactions can be defined applying specific design principles for the generation of new gesture-object pairs. In particular, to add new gesture-object pairs to a system, I suggest using two design principles proposed by Beaudouin-Lafon and Mackay [24] for visual interfaces, i.e., *reification* and *polymorphism*. The first principle, *reification*, consists in introducing new elements in the interaction, i.e., for TGI, objects and gestures. There are two strategies for introducing objects in a TGI system: using existing objects of the everyday world or building new ones, designed on purpose for the interaction. The rapid evolution of 3D printers enables an additional scenario in which new objects can be even created and introduced in the interaction scenario during system operation. Similarly, existing everyday objects can be added to the system during daily operation (cf. [146]), enlarging the number of possible interactions with minimal learning effort for the user. However, not all application domains are suitable for object reification, because having many physical objects to interact with would require a considerable space to store all the objects and to allow an easy access during the interaction with the system. Introducing gestures that can be applied to objects, the number of possible interactions can be increased consistently without increasing space requirements to allocate physical objects. Gestures have a different limitation: the user generally is not able to learn a large number of gestures, especially if they should be learned all together. A thumb rule for this limitation can rely on the Miller's magical number of 7 ± 2 [139], which depicts the maximum number of information pieces that an average human can hold in working memory. To enlarge the number of interactions without increasing the number of gestures and the number of objects, the second design principle, *polymorphism*, can be applied. Indeed, whenever a system function can be associated to different objects of the system, the same gesture can be used to perform this function on each object, e.g., playing media associated to a tagged object. This approach, called by Carrino et al. [42] functional gestures, is particularly suitable for all the interaction scenarios where the user interacts with a large number of objects (an example is provided also in Section 7.5). The *polymorphism* design principle can be applied also in the opposite manner, using different gestures to transform a generic object in a multipurpose tool. This strategy, called by Vaucelle and Ishii identity reinvention [217], allows to create new interactions exploiting the power of the verb metaphor (cf. Section 4.4). Indeed, using the same object and grasping or moving it in different ways it is possible to obtain virtually many other tools. As an example, Figure 65 in Section 7.6 shows how a generic pen can be transformed in other tools by grasping it in different manners. This polymorphism strategy is particularly suitable for mobile scenarios, where the user cannot carry many objects to interact with, or in productivity scenarios, where looking for another tool can take more time than grasping a multipurpose tool in a different manner (cf. [192]).

While experts can have an important role in the definition of the interaction of a TGI system, involving users into the definition of the tangible gesture set can bring to more usable and easy to learn interfaces. To this purpose, the *user-driven* approach for choosing tangible gestures consists in conceiving together with potential end-users the tangible gestures they would perform for the different functions that should be implemented in the system. Gesture elicitations are a popular design method to derive gesture taxonomies for particular application domains and gesture sets that can improve the user experience of a given interactive system [125, 209, 230]. The system designer can already provide the gesture elicitation participants with a set of commands or function for which the participants should suggest gestures; the designer can also ask to the user which functions they would like to have in the system. Multiple iterations can be conducted in order to find first a set of functions, then a set of gestures and finally the best associations between gestures and functions. In some cases, it is particularly useful to ask to the participants to elicit gestures in pairs for pairs of symmetric functions, in order to have symmetric gestures that are easier to remember. Having symmetric gestures allows also to automatically take into account the fifth question of Bellotti et al. [25], i.e., *Accident*, to grant that users can easily recover from errors.

It is worth noting that in most examples found in literature, the object for which gestures were elicited was already provided by the designers. Indeed, formal “object elicitations” do not exist in literature. Nevertheless, interviews, focus groups or other qualitative methods can be used to understand with which objects users would like to interact in the system. In general, because of the many aspects that can be investigated, the gesture elicitation experience should be carefully designed. Indeed, gesture elicitations often suffer from the problem of legacy bias: users often are influenced by their previous knowledge of existing systems and they often propose gestures that they already know from popular interfaces, such as smartphones. Unfortunately, these interfaces are known for poorly exploring the innate manipulative and communicative human skills. Therefore, legacy bias often limits the creativity of the participants to gesture elicitations, proposing gestures that often do not exploit the rich interaction modalities allowed by gesture interaction with objects. Morris et al. [143] recently provided guidelines to obtain unbiased gesture taxonomies from the participants to elicitation studies. Still, as shown by Hoff et al. [89], these guidelines are not always effective, and legacy bias still remains a phenomenon that is difficult to counteract.

To conclude this section, it is worth noting that the expert-driven and the user-driven methods are not exclusive and they are generally mixed: for example, the results of a gesture elicitation can be filtered by the expert according to technological or safety requirements; on the other hand, a gesture set defined by experts can be discussed with users and further refined according to their feedback.

5.3 Interaction Qualities

Whether tangible gestures are designed by experts or elicited by users, it is important to define criteria to assess the qualities of the designed interactions. Hoven et al. [214] highlighted six qualities for tangible interaction, which can be easily adapted to evaluate also the interaction quality of a TGI system:

- Integrated control: The control of the system is integrated in the physical world, thus, the user can interact directly with the object through gestures, without the need of external devices.
- Integrated representation: Feedback is also integrated in the physical world, representing the state of the system in a physical manner.
- Direct control: There is a direct mapping between user's tangible gestures and system response.
- Direct representation: System feedback directly represent the state of the system without the need of further interpretation.
- Meaningful control: The user can control the system through meaningful gestures with objects, exploiting metaphors that help the user to reflect about how the system works.
- Meaningful representation: The representation of the system state is also meaningful for the user, stimulating reflection or helping the user comparing the system with existing concepts.

Hoven et al.'s qualities for tangible interaction can be used as reference also to evaluate tangible gesture interactive systems, although they might not constitute a complete evaluation tool. Indeed, since tangible gesture interaction aims at providing richer experiences to the user through interactions that better exploits human innate skills, many other qualities could be considered to evaluate a TGI system. For example, the ability of the system to elicit stronger emotions in the user could be an additional parameter that should be considered in the evaluation of the designed interactions.

Other quality criteria, such as the five questions of Bellotti et al. [25] or the interaction concepts of Hornecker and Buur [95] can be also used as criteria for evaluating the quality of tangible gesture interaction for a particular application. However, it is worth noting that not all quality criteria must be respected in a tangible interactive system: indeed, one or more criteria could be "broken" on purpose, according to the aim of the application, for example to make the user step back and reflect [92]. Therefore, the criteria proposed in this subsection should not be used as a benchmark to compare different TGI systems (e.g., stating that the more criteria are satisfied, the better the interface), but as a tool to analyze the peculiarities of each interface, reflecting about the influence of each criterion on the overall user experience. Examples of the application of the interaction quality criteria are presented in Section 8.3.1.

5.4 Common Practices for Popular Application Domains

Hoven and Mazalek [213] identified several application domains for TGI: communication and collaborative applications, education, collaborative design, mobile applications, entertainment and gaming. As discussed in Chapter 3, analyzing examples in literature I could identify some additional domains that are worth discussing in this section: the interpersonal and emotional communication, the control of devices in different contexts and the individual and collaborative production of media content. In particular, in this section, I will discuss which types of tangible gestures seem more appropriate for each application domain, according to existing work found in literature.

In the domain of emotional design, hold and touch assume a very important role. Hold and touch are often associated to intimacy in social interactions and have a very important role in childhood [200]. While touch is important to stimulate the tactile receptors of the skin, the hold component, in particular when associated to pressure, for example while hugging, can stimulate additional mechanoreceptors facilitating the release of hormones and increasing the sense of belonging, either in relation with another person or with an object. Several artistic exhibitions make use also of continuous body movements to communicate emotions, often in association to music or lights [101].

Storytelling is a domain where tangible gestures are particularly interesting. Objects play an important role in stories [75, 198] and can be animated by children through hold + move gestures [217]. In learning and education, objects are useful for the reification of abstract concepts: by holding and moving objects it is possible to highlight relationships between different concepts embodied by the objects [166].

In work environments, where efficiency in content production is crucial, tangible gestures can be used as shortcuts that are easier to remember and quicker to perform in respect to traditional GUI or dedicated buttons [192]. The MTPen [192] adopted different hold + touch gestures to change ink-mode, while touch + move gestures while holding the pen provided additional shortcuts to system functions. In general, the introduction of hold + touch gestures to change the operating mode of a multipurpose tool allows reducing the number of interactive objects, while conserving part of the tangible affordances provided by the manipulation of a tangible tool (see also the *polymorphism* design method proposed in Section 5.2). When the work environment is set up for collaboration, tangible gestures should be designed to be easily understood by all participants to the interaction, in this case, hold + move [101] and touch + move gestures on tabletop [226] are particularly suitable, especially if the system is persistent and the effect of each gesture is clearly represented.

In gaming applications, tangible gestures obtained a great success, especially with the Nintendo WiiMote and the PlayStation PSMove controllers. Movement has a key role in gaming and the names chosen for the two controllers reflect this importance. In fact, hold + move gestures are often used for these applications and they are able to give new meanings to the multipurpose controller, according also to the different gaming contexts. As stressed by Vaucelle and Ishii [217], similarly to storytelling, these gestures are able to animate the object and to give it a new life. In contrast to work environments, in gaming applications large movements are not always avoided and fatigue could be a key element of the playful experience.

Finally, gestures are very often used to give commands and to control the behavior of digital or physical objects. Touch and touch + move gestures on remote controllers [115] or on wearable devices [14, 153, 157] are often used because of the little effort that is necessary to perform them.

5.5 Exploiting Object Affordances

Sometimes designers can find inspirations from everyday objects and can look for the gestures that are more suitable for that given object. Affordances help users to guess the interaction and simplify the work of designers to communicate to the user the available gestures in the interface. I identified several affordances exploited by existing systems: *constraints, moving parts, forms, dimensions, deformability, semantics,* and *life-like* affordances. This Section will deepen the analysis of each type of affordance.

Constraints are classic affordances exploited since Bishop's Marble Answering Machine [28] and formalized by Shaer et al.'s TAC paradigm [187]. By limiting the interaction possibility with surfaces or bindings, the users can easily guess which gesture they can perform. Tabletops are a typical constrained setup where users are likely to perform touch + move gestures on the surface around the objects or planar hold+move gestures over the tabletop [101]. *Moving parts* of the objects are able to constrain the paths of touch+move gestures [115] or hold + move gestures performed on that part [158]. The *form* of an object and in particular its ergonomics, instead, suggest to the user how to grasp (hold+touch) the object, which can be associated to different operating modalities of the object. Gestures could vary also according to the different object *dimensions*: bulky objects are difficult to hold, thus, touch + move gestures are preferable, unless they have moving parts that can be held and moved [64]. Small objects, instead, can be easily held in the hand and moved. *Deformable objects* can be distinguished in two types: objects with form memory [131] and objects without memory [125, 178]. Deformable objects often encourage the user to apply forces, both in relation to hold gestures or to touch gestures. Object movement is generally involved as an effect of the applied forces [125, 131, 178].

A particular class of object affordances is related to the *semantics* of the object. The previous knowledge of the user can be exploited to communicate through the object form the gestures that can be performed with it. The object can be an everyday object or an artifact that resemble to a well-known object or tool. This class of objects falls into the iconic category of Ullmer and Ishii[205], or the noun + verb category of Fishkin [68]: because the users know how to hold and move the object for the everyday use, they are likely to think that a similar gesture is implemented also in the digital interactive system. Particular objects are those that show anthropomorphic or zoomorphic affordances [184]. In this case, the user expects a *life-like* behavior from the object, and the designer should adopt gestures that include touch and hold components in order to obtain a greater emotional involvement of the user [200, 235].

Whenever the object is not able to communicate the possible gestures through its form, the TGI designer has to communicate to the user which gestures she can perform with the object. Typically, the designer provides explicit instructions to the user either with a tutorial or a manual. Recently, Lopes et al. suggested a new vision, called Affordance++ [133], where the object instructs the user about how to perform gestures with it. Indeed, the authors adopted a wearable system that forces the users to perform the right gestures for each object by contracting their muscles with electric stimuli.

5.6 Designing Feedback

Tangible gesture interaction can be applied to different application domains, with a large variety of gestures and objects, which makes difficult to have a complete characterization of all tangible gesture interactive systems. In particular, feedback can assume in tangible gesture interactive systems an even larger variety of forms. From the analysis of existing TGI systems, I could not find association patterns between gesture types or object affordance types and different types of feedback. However, as a general remark, since the physical world is a foundation of Tangible Interaction [214] and a criterion also for TGI (cf. Section 4.1), feedback should be delivered to the user in the physical world, better exploiting the perceptive skills of the user and their spatial awareness. Fishkin framed possible outputs of tangible interfaces according to four level of embodiment: in *full* embodiment, the system output is coupled to the system input, integrated into the object; in *nearby* embodiment, the system output is provide near the input, generally next to the object with which the user interacted; in *environmental* embodiment the output is distributed in the environment, for example through spatial audio; finally, in *distant* embodiment the output is provided elsewhere, for example in a distant screen. Generally, best output modalities are those tightly coupled to the input objects, as suggested by the interaction quality criteria presented in Section 5.3. Concerning output modalities, instead that providing feedback solely inside a screen, taking advantage only of the visual channel, tangible gesture interactive systems should take into account also the other human perceptive skills. Haptic feedback can be easily coupled to gestures that include the touch and/or the hold component, while gestures that rely only on the move component could be associated with haptic feedbacks only through wearable devices or air stimulation. Designers could explore also other sensory modalities, which are often neglected in traditional interfaces, such as spatial audio, or smells. Although alternative sensory modalities should always be explored, in particular, to support interactions that happens in the periphery of the user attention, often feedback is designed only for the visual channel. In tangible interaction, two main types of feedback can be designed for the visual channel: direct light (either emitted through LED or screens integrated in the physical world, or projected in the physical space or over the object) and physical actuation, to arrange spatially objects or to physically deform them.

Independently from the type of feedback, two main types of system behaviors are possible: either the system provides continuous feedback to the user while the tangible gesture is performed, allowing direct manipulation of a digital or physical variable, or the system responds when the user's tangible gesture is ended, considering tangible gestures as commands for the system. These two behaviors of tangible interactive systems are profoundly different and, while the first behavior is typical in tangible interaction, the second behavior is more common in gesture interaction. Although one could argue that direct manipulation could not be considered as a gesture, I believe that in tangible gesture interaction both behaviors should be considered, whenever the user performs a tangible gesture to communicate something to the machine, either expecting continuous feedback and control from the system or to perform a specific time-bound command.

5.7 Assessing and Iterating

As discussed in Section 5.1, obtaining a system that matches the needs for which it has been designed and that meets user's expectations requires several design iterations and testing. Independently from the chosen design method, extensive testing should be conducted in order to assure the quality of the final tangible interactive system. While the interaction qualities discussed in Section 5.3 can be designed a priori, there is a number of other parameters that should be assessed a posteriori with proper testing. Typical quantitative evaluations for tangible gesture interactive systems are related to the classifier abilities to recognize users' gesture (see Section 6.3 for more details). In particular, typical quantitative parameters that can be measured are the F1-score and the overall recognition accuracy.

As for interaction qualities, the parameters that should be evaluated vary according to the application domain. For example, in work environments and especially for media production, a tangible gesture interactive system should be as much efficient as possible, speeding up the work of the users. In this case, the gesture recognition accuracy, ideally, should be as close as possible to 100%, with no miss-classification cases and few unspotted gestures. For other application domains, such as emotional applications, the recognition accuracy of the system might be not crucial, and misclassified gestures could not affect the perceived experience for the system (see Section 7.3 for an example). In this case, the system designer could be interested in measuring the user engagement, for example, counting the number of interactions over a long period of time (one week to several months) and measuring the user experience and the emotional involvement while interacting with the system. To this purpose, the designer could use standard questionnaires such as the System Usability Scale (SUS) [31] to measure usability (and learnability) and specific questionnaires for assessing the positive/negative activation and valence (e.g., the PANAVA-KS [181]) to determine the emotional experience generated by the system. Qualitative analysis allows understanding better the user's perception of the system: the think-aloud protocol is often used during system testing to this purpose. To understand better the emotional reception of the system one could also analyze the uses' facial expression while testing the system, with either computerized facial expression recognition or manual annotation of the recorded videos. Finally, semi structured interviews and questionnaires could help investigating specific parameters of the system, in order to individuate unspotted problems and potential solutions.

5.8 Conclusion

This section presented the designing dimension of the framework, providing high-level guidance to design tangible gesture interactive systems, describing the different phases of the design process, which are mostly similar to design phases of other interactive systems. Nevertheless, the peculiar signs used in TGI for human-computer communication, i.e., gesture-object pairs, make the interaction design for TGI systems different from previous disciplines, because the designer has to take into account the different forms that objects and gestures can assume (which have been described in Chapter 4) and the impact of these forms on the final system. Following these considerations, I suggested two different design methods to define the tangible gesture set of a TGI system, which could be either designed by experts according to best practices and existing knowledge in the application domain, or suggested by users through gesture elicitation studies. In Section 7.2, I will present the result of the application of both design methods for designing two different systems for the same application domain, i.e., interacting with the infotainment system of the car. For the expert-driven method, I suggested two design principles for the generation of new tangible gestures: reification and polymorphism. It is worth noting that polymorphism is particularly interesting for TGI, since it exploits the dualism gesture-object to assign different meanings (i.e., different system functions) to the same object or to the same gesture, by combining it with different gestures, or objects, respectively.

With the purpose to help designers choosing among the large palette of possible tangible gestures, Section 5.4 and 5.5 resumed the best practices found in literature for the design of tangible gestures, with a particular reference to the taxonomy of tangible gestures proposed in Chapter 4. In particular, they presented, respectively, the common gesture choice for different application domains and the typical object affordances that can be exploited for tangible gesture design. Concerning feedback for TGI systems, I could not find a particular correlation between the different tangible gesture types framed in Chapter 4 and the different types of feedback. However, the analysis of different system behaviors presented in Section 5.6, allowed to evidence the particular dualism of TGI interactions, which included direct manipulation of objects (in which the output is continuously updated during gesture execution) and discrete gestures (in which feedback is provided only at the end of the gesture). Finally, I discussed some evaluation criteria and methods to assess the quality of the interaction in a TGI system, either a priori, before the design of the system (Section 5.3) or a posteriori (Section 5.7). The next section will provide more detailed guidelines and technical details on how to recognize tangible gestures and how to build a tangible gesture interactive system.

6 Building Tangible Gesture Interactive Systems

“ Sempre la pratica dev'essere edificata sopra la buona teorica.”

Leonardo da Vinci, “Trattato della Pittura, Parte Seconda/77”.

This chapter presents the building dimension of the Tangible Gesture Interactive Framework. In particular, it discusses four possible approaches for integrating sensors for tangible gesture recognition, it describes the hardware toolkits that can be used to recognize tangible gestures and typical algorithms that can be used to recognize tangible gestures.

6.1 Technological Approaches for Sensors

Tangible gestures can take various forms: in Chapter 4, I identified three main components (move, hold and touch) that can be combined and to which the designer can associate further attributes (e.g., pressure, amount, speed, etc.) as additional degrees of freedom. Thus, selecting a subset of gesture types, i.e., a limited number of component and property combinations, can be useful to restrain the sensors needed to detect and recognize them. In Section 6.1, I will discuss only the different hardware sensors that can be adopted in a TGI system, the hardware platforms and toolkits to support the development will be presented in Section 6.2 while the software (e.g., algorithms and techniques) required to segment and classify gestures will be discussed in Section 6.3.

Each year, new technologies for recognizing gestures with objects are proposed. Recently, Sato et al., demonstrated a new capacitive sensing technique to make every object and also our body touch and grasp sensitive [180]. As an opposite approach, Klompmaker et al., used depth cameras in the environment to detect finger touches on arbitrary surfaces, holding, moving and releasing objects[118].

Finding the most appropriate technologies for tangible gesture recognition is not trivial and a survey of existing ones is out of the scope of this thesis. However, I propose to frame recognition systems for TGI into three main approaches: (1) embedded and embodied; (2) wearable; and (3) environmental. These approaches differ not only in the emplacement of sensors, but also in the point of view of the interaction: respectively, from the object, from the user and from an external point of view. Moreover, advantages and disadvantages exist for each approach: an object equipped with sensors can generally observe only the tangible gestures related to it, while using a wearable system generally it is possible to recognize gestures that the user performs with whatever object. In the environmental approach, instead, the same sensors can observe gestures made by all users with all objects. Indeed, an approach can be preferred to another (or possibly combined), depending on the interaction scenario and power/space requirements. A comparison of the different approaches according to different design parameters is proposed in Section 6.1.4. Section 6.1.4 presents also the possible advantages of mixing different technological approaches into the same system, introducing the fourth technological approach, i.e., hybrid.

As tangible gesture interaction starts permeating our common activities and our everyday environments, gesture spotting becomes a challenging task for the system designer. Distinguishing gestures communicated to the systems from common everyday activities is critical to avoid unintended responses from the system. The presence of an object to interact with generally facilitates gesture spotting, because the gesture begins only when the user gets in contact with or in proximity to that object. Still, a more fine-grained segmentation is needed, because several gestures can be performed together without leaving the object from the hand. Most gestures can be delimited by changes in one component (presence or absence of move, hold or touch) or in the value of an attribute. However, there are some examples [171] in which this event is difficult to recognize and an external trigger is needed. This trigger can be a vocal command, pressing a button or an additional gesture. In order to recognize the different tangible gestures, four main tasks have been identified: (i) the identification of the object with which the user is interacting; (ii) the recognition of object or hand movements; (iii) the recognition of the way the user is grasping the object, i.e., hold + touch gestures; and (iv) the recognition of touch + move gestures on surfaces. Finally, the designer should also include proper feedback in order to acknowledge the user's commands.

6.1.1 Embedded and Embodied

The *embedded and embodied* is probably the most common approach for TGI. Enhancing objects with onboard technology is a well-known practice in the field of tangible interaction. As shown by Atia et al.[15], even simple objects can be augmented with gesture recognition capabilities by just attaching a small module to it. Moreover, the identification of objects with which the user is interacting is generally a trivial task, since each object is aware of the user interactions (conversely, user authentication or recognition could be a complex task). Inertial sensors are sufficient to detect most hold + move tangible gestures [15, 64]. Touch sensing, both static and dynamic, can be achieved with different techniques: capacitive sensors [180, 192], pressure sensors, and cameras that analyze light conveyed from the surface of the object[227].

6.1.2 Wearable

The *wearable* approach exploits technologies embedded in worn accessories, especially those very close to the core of the interaction, e.g., the hand. The wearable approach overcomes most problems of occlusion typical of cameras placed in the environment and allows the interaction with many common objects, which do not need embedded sensors. In fact, recognizing objects can be done sticking an RFID tag on them and integrating an RFID reader in a glove [198] or in a wrist-band [63]. The detection of hold + move gestures is easy when the movement of the forearm is rigidly coupled to the movement of the object. In this case, inertial sensors allow detecting most of possible gestures with the object [63, 142]. However, detecting the movement of the object within the hand, like pen rolling, is more difficult. The movement of the object could be inferred by inertial sensors placed on the fingers, for example in a ring [114], or analyzing the movement of tendons in the wrist [142]. The analysis of tendons is an interesting approach also to detect grasp postures and touch gestures.

6.1.3 Environmental

The *environmental* approach offers an external point of view for the analysis of tangible gestures. This point of view is placed in the environment, which allows discovering important properties of tangible gestures. Seen from the environment, hold+touch gestures are obviously static; touch+move gestures imply a movement of the user hand or fingers relative to the object, which, instead, is generally static; hold+move gestures involve a movement of both the object and the user's hand with respect to the environment. Thus, depending on the gestures type, a vision based recognition system should focus on the tracking of the hand, of the fingers, or of the object. Three main trends can be identified for the environmental approach. Tabletops [101, 226] are very common in tangible interaction and restrain the interaction space to a surface, which generally simplifies the recognition tasks. Systems with fixed RGB [27] and/or 3D [118] cameras in the environment are also possible, even if they can suffer from occlusion problems. Finally, as shown by Li et al. [128], gestures could be recognized by mobile cameras in the environment, which could be integrated in serving robots that follow the user in the environment.

6.1.4 Hybrid Systems and Parameters for System Design[16]

A TGI system can be built integrating technology in the physical world according to multiple approaches, mixing the embedded and embodied, the wearable and the environmental approach. Whenever two or more approaches are adopted, in these cases, I will call the resulting system *hybrid*. Data coming from the sensors integrated in the physical world according to the different approaches can be collected by the same system and exploited altogether to recognize tangible gestures. Alternatively, a whole system can be built according to each approach, and the resources offered by each system (such as gesture recognition capabilities, but also interaction modalities and peculiarities) can be shared and opportunely used to obtain a system that provide the best possible experience to the user. Rhodes et al. already showed the potential advantages of mixing wearable and environmental systems [172] in the context of networked devices, with a particular focus on the complementarity of some aspects offered by the two approaches. Indeed, Rhodes et al. showed that while the wearable paradigm supports *privacy* and *personalization*, the environmental solution can bring *localized information*, *localized control* and *resource management*. Carrino et al. [43] extended these aspects introducing the concept of consistency and focusing its analysis on the improvement of gesture recognition systems through classifier fusion. The purpose of integrating two or more heterogeneous systems is exploiting their complementary characteristics, in order to obtain an opportunistic synergistic system that offers advantages over the single systems. This synergy could be applied only

[16] This section of the framework has been pushed in 2013 in [10] in collaboration with Stefano Carrino, Francesco Carrino and Maurizio Caon. It is worth mentioning that in this thesis I distinguish between the embedded and embodied approach and the environmental approach, while in the article [10] sensors embedded in the objects are considered as environmental sensors.

when needed, whereas in some occasion it would be preferable to use only one system. To this purpose, I propose the following definition of opportunistic synergy [10]: *an opportunistic multi-component system is synergistic if all the components contribute to the overall system performances, but the lack of some of them does not compromise the proper functioning of the whole system. A system with multiple components should perform better than its individual components. Performances are measured in terms of interaction, usability or accuracy.*

In particular, in order to evaluate the advantages of combining systems designed with different technological approaches, I have individuated eight design parameters [10]:

Interaction area is the space in which the user interactions and commands are sensed by the system. While wearable systems allow the user to virtually interact everywhere, as long as they are worn, embedded and embodied systems limit the interaction area to the surroundings of the object where sensors are embedded, and environmental systems are generally able to track up to a whole room, with possible limits of occlusion.

Personalization is the capacity of the system to provide and maintain personalized interactions for the users. While wearable devices offer the greatest support for personalization, environmental systems are generally used by different people and are more difficult to customize. Systems embedded in objects could allow for personalization especially when the object is personal, i.e., used only by one user.

Consistency is the capacity to improve the system thanks to the prolonged, continuous interaction between the human and a computer. As for personalization, wearable systems and personal smart objects could improve their performance by learning from the specific user.

Private interaction and intimate interfaces are "*discreet interfaces that allow control of mobile devices through subtlety gestures in order to gain social acceptance*" [51]. Wearable systems and personal embedded and embodied devices offer a higher level of privacy in respect to environmental systems.

Localized information is the system feature that specifies how to access the information in a specific location. While in wearable systems information is provided directly on the user and is always at user’s glance, in environmental system it could be distributed in the environment or localized in a specific region, while in embedded and embodied systems is localized in the object. While having all the information at user’s disposal could be interesting for some applications, in other cases, distributing the information in the location of interest could be preferable, avoiding overwhelming the user with unnecessary information.

Localized control is the system feature that specifies how and where to provide information and commands to the system. As for the localized information, the control is always available to the user in wearable systems, while in the environmental systems could be localized in a region or in an entire room and in the embedded and embodied systems is localized in specific objects. Again, a localized control could be preferable in some application to avoid unintended commands or to differentiate different commands according to the location or the object.

Resource availability is strictly linked to the current technologies adopted for the interaction, e.g., processing power, energy, etc. Wearable systems and embedded and embodied systems generally rely on batteries for energy and thus have limited processing power, while environmental systems have often higher resources.

Resource management is the system capability to handle efficiently the different available resources. Environmental systems generally are able to manage a larger network of sensors and resources, while wearable and embedded systems have a more limited scope, although the efficient management of these fewer resources is key to save processing power and energy resources.

The design of tangible interactive systems obtained mixing different technological approaches should take into consideration all these parameters in order to evaluate which approach is better for a particular application and whether combining multiple approaches could enrich the interaction for the user and the efficiency and accuracy of the system. Further details about the increase of the recognition accuracy of hybrid systems is presented in section 6.3.3.

6.2 Hardware Toolkits and Parameters for Platform Choice

Building tangible gesture interactive systems requires hardware that can be easily integrated in the physical world, through low power microcontrollers able to retrieve data from different sensors, to provide output for the user and to connect to networks for exchanging information with the external world. Ideally, these systems should be powered by a battery and should be as small as possible, in order to disappear to human eyes. At the same time, in order to speed up the development and allow for rapid prototyping, trying out different sensors and different configurations, these systems should be easily configurable and hackable with minimal development effort.

During my thesis, I focused my attention on two technological approaches, the wearable approach and the embedded and embodied approach. For practical reasons, the environmental approach has not been explored during this thesis. At the beginning of this thesis, I looked for a platform to develop a smart watch, with enough computational resources to perform tangible gesture recognition inside the platform. The chosen platform was the Amadeus APF27 board, an embedded development platform based on ARM926 processor and running Linux, which could ensure connectivity with I2C sensors and enough processing power to elaborate data onboard. Some inertial sensors had been successfully tested on the platform, but it required a consistent effort and a good knowledge of the sensor registers, since no compatible sensor library was available for this platform. Moreover, realizing a small prototype to be tested in a real case scenario was particularly difficult because of the size of the board.

For these reasons, in the following months I decided to develop new prototypes using the Arduino platform, which allowed seamless and effortless integration, thanks to the numerous libraries and the easiness of configuration of the different sensors. Although the computational power of the Arduino board was very limited, it allowed designing several different systems, with minimal effort. Because of the limited computational power, I needed to send sensor data to a more powerful platform, such a PC, in order to complete the gesture recognition task. Nevertheless, the Arduino platform allowed the rapid development of ready to use proof of concepts with minimal effort, often with a high level of integration in the objects, which made technology disappear to the users' sight. As a declination of the Arduino platform, the Lilypad Arduino allowed a seamless integration in wearable systems, making technology disappear in the fabric of everyday clothes (cf. Section 7.4.5).

Obviously, many other platforms that are available on the market, thus the choice of the right hardware platform for developing an embedded and embodied or a wearable tangible gesture interactive system is not easy. Moreover, thanks to technological advances new platforms are released each year, increasing for the designers the possibilities that will be available out of the box. For these reasons, any platform that would be recommendable today could easily become outdated in the next months. Based on the knowledge acquired during my thesis experience, I summarized a list of features that should be traded off to evaluate which is the best hardware platform that could satisfy the specific needs of the TGI designer:

Computational resources: Each embedded platform offers a different microcontroller with different computational capabilities, depending on the microcontroller architecture and speed, as well as on the available memory at disposal of the microcontroller. These specifications range nowadays from the 8-bit RISC microcontroller operating at 20 MHz with only 2KB of RAM and 32KB of flash memory (Arduino Uno), to Quad-core 64-bit System on Chip operating at 2.4 GHz, with 6GB of RAM and 256GB of flash memory (Snapdragon 821 on most recent smartphones). While in many cases even the less powerful microcontroller, e.g., the Arduino Uno, is able to acquire data from more than one sensor, to perform a basic data elaboration and to enable actuators for user feedback, running complex algorithms on large amount of data (for example whenever computer vision is needed) requires a more powerful platform, for example a Raspberry Pi. Moreover, some platforms, due to limited computational capabilities, offer limited support to advanced programming techniques, such as multi-threading and event-based programming.

Power requirements: While most embedded platforms can be powered through a USB port or through an AC/DC adapter with a power plug, in many cases the designers need to develop wire-free systems that can be easily moved by the user. In these cases, the system should be able to operate on battery power, which should last at least for one session of interactions (up to one day of operation). Moreover, the battery should be easily rechargeable or replaceable with minimal effort by the user.

Extensibility and connectivity: the platform should allow to retrieve data from several sensors, which could be either digital or analog, and to be able to drive different actuators. In particular, different standards are available for digital sensors, the most common being the I2C and SPI buses and different serial connections (UART, RS232, USB, etc.). To connect analog sensors, an Analog-to-Digital Converter (ADC) must be present on the board. Often, the extensibility of the development platform can be roughly assessed by the number of I/O pins. Finally, in order to communicate with other devices, sometimes network connectivity is needed. Many boards are already equipped with Ethernet, WiFi or Bluetooth connectivity, while other boards allow extending connectivity through shields or through dedicated devices to be connected to I/O pins.

Easiness of physical integration: Physical dimensions (of the board and the required battery) are often the biggest constraints to embed TGI systems in physical objects. Nevertheless, bigger development boards have generally dedicated headers to connect sensors, speeding up the rapid prototyping in respect to boards that require soldering wires. The availability of electronic schematics to print your own board with all the needed sensors and your custom layout is another crucial parameter, especially when designers aim at developing a well-integrated commercial product.

Development speed: The availability of precompiled libraries for the chosen platform and of user-friendly interfaces to build the software could speed up the development of the software logics to control sensors and devices as well as of the algorithms for recognizing tangible gestures.

Level of control and hackability: Some platforms hide many low-level functions in already-provided libraries. Similarly, often some platforms offer plug-and-play connectivity for sensors through proprietary connections and buses. Libraries and connections that are not open-source and that cannot be customized by the TGI designer could limit the possibilities of development and the extensibility of the board to non-standardized sensors.

Community support: Depending on the diffusion of the platform worldwide, it will be possible to find existing examples and tutorials on the net that could ease and speed-up the development of the TGI system. In some cases, community and forums can provide direct support for the project and hints to solve common problems. Newer platforms or less diffused one could offer limited support in this sense.

6.3 Software for Gesture Recognition: Techniques and Algorithms

Four main different tasks have been individuated in Section 6.1 for the recognition of tangible gestures. In this Section, I will not deal with the first task, i.e., the identification of objects, which is trivial when objects are identified through RFID (wearable approach), or when objects are already equipped with technology (embedded and embodied approach), while it becomes a complex task when the recognition is performed through computer vision (environmental approach). As stated before, in my thesis I did not investigate computer vision techniques for the recognition of tangible gestures. The other three tasks deal with the recognition of gestures, either implying movements, i.e., hands or objects movements (move, hold+move) and finger movements on surfaces (touch+move), or static finger touches on objects or surfaces (hold+touch, touch). As discussed in Section 5.6, two main system behaviors can be designed: either the system provides continuous feedback to the users while they are performing the gesture (e.g., for direct manipulation), or the system replies to the users only when the tangible gesture ends. The first behavior implies a particular recognition system that is able to couple continuously the user input to the system output. This requires continuous evaluation of users' tangible gestures, and, in particular, an estimation of the parameters associated to the output with a transfer function. Different techniques can be implemented to obtain such behavior, but will not be discussed in depth in this thesis. Instead, I will deepen the discussion on systems that do not require continuous evaluation, thus, on systems that perform the recognition at the end of the users' gesture. In particular, I will deal with three different topics: (i) gesture segmentation, (ii) gesture classification, and (iii) segmentation and classification fusion (for hybrid approaches). Feature extraction (as discussed by

Pavlovic et al. [155]) will not be discussed in this thesis since it could vary largely according to the application domain and the sensors used to recognize gestures.

6.3.1 Gesture segmentation

As evidenced in Section 2.3.3.3 for gesture interaction and as discussed in Section 4.3.2, also in tangible gesture interaction the delimitation of gestures in time is a key issue that should be addressed already during the design phase, in order to facilitate the spotting of gestures and the accuracy of the recognition. As discussed in Section 4.3.2, a gesture is delimited in time by a change in one component or attribute. In some cases, tangible gesture interaction offers an implicit support for gesture segmentation. For example, a hold+move gesture could be easily delimited through two events: the user take the object in the hand (gesture start) and the user release the object from the hand (gesture end). Similarly, a touch+move gesture could be easily isolated detecting when the user is touching the surface of the object. These considerations allow reducing the system attention to a shorter interval. However, in some other cases, a more accurate segmentation could be required. For example, the designer could be interested in isolating the hold+move gesture from the preparation and closure phase of the gesture, which include picking up the object and reaching the desired start position as well as leaving the object when the gesture end. Similarly, the designer may need to separate subsequent gestures that the user could perform without releasing the object from the hand. Two strategies are possible: an explicit segmentation, operated by the user through an additional device or interaction modality, or an implicit segmentation, operated by the system through an automatic analysis of the signal. While the first strategy could be more accurate than the second one (since the users knows when they want to communicate something to the machine), it requires additional cognitive and motor resources from the users.

6.3.2 Gesture Classification

As for gesture interaction [155], the purpose of gesture classification in TGI is associating to an observed tangible gesture one of the gesture classes that have been conceived for the system. The advantage offered by tangible gesture interaction is the possibility to enlarge the gesture set without increasing the difficulty to recognize gestures: adding more objects to the system will allow increasing the gesture-object pairs without defining new gestures. Still, the recognition of the gesture is a key point that should be carefully addressed to build a tangible gesture interactive system. As for gesture interaction, the observed gesture can be represented through a vector of data, which can be either the raw data collected by the sensors or features extracted from the raw data. In particular, two types of gestures should be considered: static postures, such as hold+touch gestures, and dynamic gestures, such as hold+move and touch+move gestures. While in static postures, the variation of the data over time, between the start and the end of the gesture is not important, in dynamic gestures, the temporal variation is particularly important and should be taken into account by the gesture classifier. Often, data related to dynamic gestures are considered as time series.

Two main types of gesture classifiers can be identified: those based on *heuristics* and those based on *machine learning*. Gesture classifiers based on heuristics exploit algorithms developed on purpose by the system designer to recognize gestures based on predefined criteria. Most simple heuristics are based on thresholds, which are determined by the designer after inspection of sensor data. Gesture classifiers based on machine learning (or supervised learning) rely on functions that are automatically generated by the system and that are able to determine to which class an observed gesture belongs. In order to generate these functions, the designer must provide to the classifier sufficient labeled examples for each gesture class. To ensure the genericity of the system, these gesture examples should be performed by different users, providing several repetitions for each gesture class. In this manner, the generated functions should be able to cope with typical variations that can occur across gesture executions of the different users. An alternative approach granted by machine learning classifiers is training the system only on one person, i.e., the one that will use the system. This is particularly interesting for all systems designed following the wearable approach, as well as for the systems embedded in personal objects. The possibility to train the system only on the user not only allows increasing the accuracy of the classifier, but allows also the user to define his or her own set of gestures, without the need to learn gestures chosen by the designer. Several classification algorithms based on machine learning exist. Among them, Hidden Markov Models, Support Vector Machines and Linear Discriminant Analysis are often used for gesture classification. The choice of the different algorithms to be used depends on the specific application and in particular on the data collected by the sensors. Moreover, most algorithms allow tuning several parameters to optimize the recognition accuracy. This operation requires additional data (independent from the training set) to test the performances of the classifier with the different parameters.

It is worth noting that two type of outputs are possible for machine learning classifiers: hard classifiers provide crisp results, assigning an observed gesture to a class or to another one; fuzzy classifiers are able to provide soft results, generating as output the likelihood (a probability that ranges from 0 to 1) for each class that the observed gesture belongs to that class (the chosen class will be that with the highest likelihood). This additional information is particularly interesting when comparing the results of different classifiers and, in particular, for the problem of classifier fusion, which is discussed in the following Section.

6.3.3 Classification and segmentation fusion[17]

As discussed in Section 6.1.4, in some cases it is interesting to benefit of tangible gesture interactive systems developed with different approaches and to merge them into a single synergistic system that exploits the best capabilities of each sub-system. In particular, each sub-system is equipped with different technologies for spotting tangible gestures and for classifying them. Therefore, an intelligent strategy for merging the classification and segmentation results of each sub-system should be found.

In order to merge the result of different classifiers, an approach based on late fusion has been investigated [10]. Since the fusion is applied only at the last stage of the classifiers, the synergistic system would not be affected by changes in the sub-system classifiers. In the context of my research, I investigated and compared 10 different methodologies to weight and merge the results of each classifier are presented in this section. While some methodologies were taken from literature others have been conceived and assessed in the context of this research [10]. These methods can be applied only to classifiers with soft outputs, while some methods cannot be applied to more than two classifiers. Most of these methods rely on a weighting process that is based on the confusion matrices, which gives an a priori knowledge of the ability of each classifier to recognize correctly a given gesture (and the probability to miss-classify it with another gesture class). Confusion matrices of the classifiers can be estimated in a cross-validation phase (for example, performing a k-fold cross-validation on the training set). For brevity purposes, the proposed methods will not be reported in this thesis. The reader could refer the published article [10] for further details on the proposed fusion methods.

For the realization of a synergistic paradigm (cf. Section 6.1.4), a critical step is also the processing of the segmentation signals. Mainly for two reasons: first, from an interaction point of view the segmentation can have an impact on the cognitive load on the user; second, having sub-systems working asynchronously, the fusion system can receive gestures that are delayed and it can deal with missing signals. The fusion classifier should deal with the possible delay between the outputs of the different classifiers by fixing a maximum delay d that should be accepted to consider two asynchronous sub-system outputs as belonging to the same gesture. When the delay between the outputs of the two classifiers is greater than d, gestures are treated separately and results are not merged. This approach

[17] This section of the framework has been pushed in 2013 in [10] in collaboration with Stefano Carrino, Francesco Carrino and Maurizio Caon.

implies that the user should wait at least the time *d* between a gesture and the following one. In a synergistic system, three different strategies can be implemented to merge the segmentation decisions coming from two or more sub-systems. In the case of two classifiers, the OR strategy considers as valid results the contribution of both single and coupled classifiers (as the logic OR). The AND strategy takes into account only gestures simultaneously detected by the two classifiers. The ADAPTIVE strategy opportunistically switches between the OR and the AND strategies, taking into account the a-priori probability that the classifiers could have erroneously spotted (false positive) or missed (false negative) a gesture. This a priori probability can be calculated from the confusion matrices of the single classifiers.

6.4 Conclusion

This Section provided guidance on how to build tangible gesture interactive systems, both from a hardware and software perspective. In particular, it presented three different approaches (embedded and embodied, wearable, and environmental) for integrating technology respectively in the objects to interact with, on the user, and in the environment, as well as a hybrid approach for mixing systems of different nature. Design criteria are provided to evaluate the advantages of the different approaches and what can be obtained by mixing them together. Moreover, this section presented criteria for choosing the hardware platforms for embedding tangible gesture interactive systems in the objects or on the user body, as well as a generic overview on software techniques for segmenting and recognizing tangible gestures. Examples of implementations of the embedded and embodied approach, of the wearable approach as well as of the hybrid approach are presented in the following chapter. For practical reasons, the environmental approach and computer vision techniques for recognizing tangible gestures have not been explored in this thesis.

Part III – Framework Applications and Validation

7 Framework Application: Seven TGI Systems

" La sapienza è figliola della sperienzia."
Leonardo da Vinci, "Aforismi, novelle e profezie".

This chapter describes the systems developed during this thesis as practical applications of the Tangible Gesture Interactive Framework presented in the previous chapters. In particular, it presents four systems developed for interacting with the infotainment system of the car, two systems for emotional and interpersonal communication and one system for the interaction in the smart home.

7.1 Introduction to TGIF Applications

In this chapter, I present seven different systems designed and developed during my thesis to explore the different aspects of the Tangible Gesture Interaction Framework. The seven systems encompass three different application domains, two different design methods, three different technological approaches and gesture recognition algorithms based either on machine learning techniques or on heuristics. The seven systems allowed to explore several different tangible gestures, obtained combining the move, hold and touch component and additional gesture attributes. Gestures were performed either on a specific object that provided particular affordances or on multiple objects, with generic forms. At the beginning of each subsection, I will present the aspects that each system allowed to explore in the context of TGIF. A summary of the aspects of the framework explored through the different systems is presented in Table 2. The design space of TGIF explored through the systems presented in Chapter 7 is depicted in Figure A10 in Appendix A.

Table 2. TGIF design space explored through the application examples of Section 7. W=Wearable, EE= Embodied and Embedded. M= Move, H=hold, T= Touch.

System	Application	Gesture Design Approach	Object (Affordance)	Gesture	Syntax			Semantics	Tech Approach
					M	H	T		
WheelSense V1	Control	Expert-driven (Hold constraint)	Steering wheel (form)	Hold+touch&Pressure | Hold+touch+move	x	x	x	Embodied metaphors (partially)	W
WheelSense V2	Control	Expert-driven (Hold constraint)	Steering wheel (form)	Hold+touch&Pressure | Hold+touch+move	x	x	x	Embodied metaphors	EE
WheelSense V3	Control	Expert-driven (Hold constraint)	Steering wheel (form)	Hold+touch&Pressure | Hold+touch+move	x	x	x	Embodied metaphors	Hybrid (W+EE)
WheelSense V4	Control	User-Driven (gesture elicitation)	Steering wheel (form)	Touch | Touch+move	x		x	None (User-designed, spatial reference)	EE
ADA Lamp	Affective communication	Expert-driven (common practices)	Lamp (anthropomorphic)	Touch | Touch+move | Hold + touch | Hold+move	x	x	x	Noun+Verb	EE
Hugginess	Affective communication	Expert-driven (common practices)	Human Body (anthropomorphic)	Hold+touch&Pressure		x	x	Full	W/EE
Smart Watch	Control	Mixed: Expert-Driven + User observation	Different small objects (semantics)	Hold | Hold + move (Hold+touch&Pressure | Hold+touch | Hold+touch+move)[18]	x	x		Noun	W

[18] As additional explorations during feasibility studies

7.2 WheelSense

NOTE: This section presents different studies conducted in collaboration with my colleagues Francesco Carrino, Stefano Carrino and Maurizio Caon, which have been published in [7, 9, 10]. These studies have been partially integrated also in their respective PhD Thesis.

7.2.1 Introduction and Field Requirements

Controlling the infotainment system of the car can be a distracting secondary task. In this context, improving the user experience of the driver while maintaining high his or her attention on the primary task, i.e., driving, is a key requirement. Indeed, these systems should require little cognitive attention and should interfere minimally with the perceptive skills (mostly vision and haptics) and motor skills (hands and feet) used for driving. The system should be also robust and although errors can be tolerated, they can generate distraction. As shown in Section 3.3, gestures can be a safe interaction means with the infotainment system since they can decrease the visual demand and the cognitive load to interact with the IVIS while driving. The WheelSense project aims at investigating gestures performed on the surface of the steering wheel to interact with the In-Vehicle Infotainment System (IVIS). In this project, four different systems (WheelSense V1 through V4) have been realized, each exploring different aspects of TGIF: three different technological approaches have been investigated to recognize tangible gestures with the steering wheel: sensors embedded in the steering wheel (embedded and embodied approach), sensors worn on the user body (wearable approach) and a hybrid approach that combines the two previous approaches. The first three versions of the WheelSense system aimed at meeting a particular safety requirement: gestures must be performed while holding firmly the steering wheel with both hands in the position suggested by the Swiss driving school manual [103]. This requirement imposes a constraint on the presence of the hold component in the gestures performed on the steering wheel. The last system, instead, has been developed following a user-driven approach, exploiting a gesture elicitation study to choose the gesture vocabulary. In this case, the aim of the system was to obtain gestures easier to learn and to perform for the driver, with the purpose of increasing the user experience, loosening the previous safety requirement of holding the steering wheel with both hands.

7.2.2 WheelSense V1 (Wearable)[19]

7.2.2.1 TGIF aspects explored with this system

This system allowed exploring the use of tangible gestures for the *control* application domain, specifically, for controlling the infotainment system of the car. The interaction was designed using the *expert-driven* method, with a *reification* of new gestures performed on the steering wheel. Following field requirements, all gestures were designed including the *hold* component. Gestures were designed partially exploiting *embodied metaphors* and the *form* affordances provided by the steering wheel. Moreover, they were chosen according to the recognition performances allowed by the selected *wearable* technology for recognizing gestures (electromyography). Software-side, *machine learning* has been used to recognize gestures. The design space of TGIF explored through this system is represented in Figure A2 in Appendix A.

7.2.2.2 Design Rationale and Methods

This system has been conceived with the purpose of customizing and enhancing the interaction with the infotainment system of a car, without modifying the existing interface of the car, which can be costly and time consuming. The system has been designed following an expert-driven design method, with the purpose of obtaining a wearable system that is able to recognize gestures performed while driving. As discussed in Section 6.1.4, a wearable system can help the personalization possibilities without modifying the interface of the infotainment system of the car, which is an optimal solution for a car-sharing scenario. Following the insights of Costanza et al. [51], ElectroMyoGraphy (EMG) has been chosen as sensing technique for the wearable approach. This technical choice has influenced the gesture design phase, since some gestures that imply little muscular activation can be more difficult to detect than other gestures. A preliminary inspection of EMG raw signals allowed to detect which hand gestures could be easily recognized by the system. Moreover, the additional constraint of driver safety was taken into account for the gesture designs: gestures should be performed by the user while holding the steering wheel with both hands. Finally, the user's cognitive load for interacting with the system should remain low, since the user's attention needs to be focused on the primary task, i.e., driving.

[19] This work has been published in [41] in 2012, in collaboration with Francesco Carrino, Stefano Carrino and Maurizio Caon

7.2.2.3 Interaction Design: Tangible Gesture Syntax and Semantics

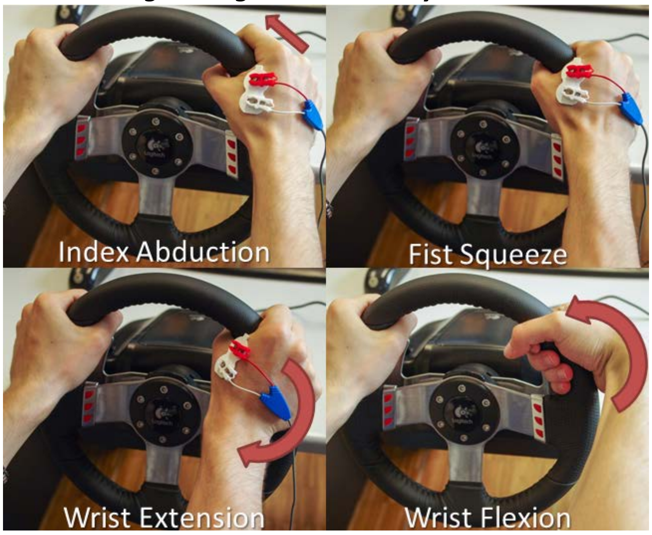

Figure 26. The four gestures designed for the wearable system.

Four tangible gestures have been designed to interact with the infotainment system. In all the four gesture-object pairs, the object is obviously the steering wheel (a Logitech G27 steering wheel, which has not been modified). Because of the safety constraint, both hands need to remain on the steering wheel while performing gestures. For this reason, the hold component was present in all the four tangible gestures. The four gestures are shown in Figure 26. The first gesture is an index abduction to start music and it is performed with a subtle movement of the index finger while holding the steering wheel. This gesture has been chosen because it can be easily recognized by the system and was economical to perform, even if it has no direct semantic reference to the associated function in the system. The second gesture is a fist squeeze to stop music and implies an additional pressure attribute over the hold component. Although there is no direct association between squeezing and stopping music, it exploits an embodied metaphor [19] to link the hand closure with the IVIS closure. Evidence about the intuitiveness of such interaction is shown also in Section 7.5. The third and the fourth gesture are performed by flexing and extending the wrist to drag the fingers around the steering wheel, upward and backward, in order to go to the next or previous song, respectively. Therefore, besides the hold component, an additional touch+move gesture is performed. Again, these gestures are not directly mapped to the respective commands, but the opposite upward/downward movement can be easily to associated the next/previous song through embodied metaphors [19] (Figure 26). It is worth noting that the steering wheel, the object on which gestures are performed, offers enough physical affordances to support the gesture execution without particular effort or pain for the user. The system was designed to work with the inherent audio feedback from the infotainment system. However, for the preliminary evaluation of the system (see Section 7.2.2.5) feedback has not been used.

7.2.2.4 Implementation

The system has been conceived and implemented following the wearable approach. The chosen technology for recognizing gestures, i.e., ElectroMyoGraphy, required sticking electrodes on the user's forearm skin to sense the electrical activity of four selected muscles (*First Dorsal Interosseous*, *Flexor Carpi Ulnaris*, *Palmaris Longus,* and *Extensor Digitorum)*, which are activated when the four selected gestures are performed (see Figure 27).

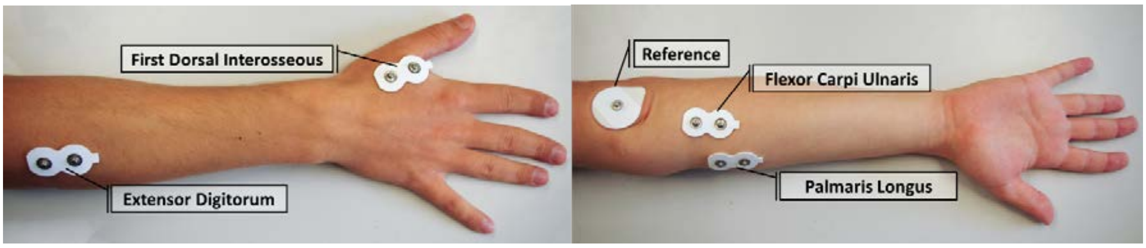

Figure 27. The four electrodes applied to the selected muscles, plus an additional electrode to be used as reference for the signal.

The electrodes are connected to a Bluetooth signal amplifier, the Marq-Medical MQ16, which sends data to a PC for further elaboration. The system is a preliminary prototype, which cannot be used yet in a real-case scenario. Indeed, this system requires that an expert or a trained person individuate the position of the muscles to be sensed, in order to stick the electrodes on the right positions on the forearm. Nevertheless, technological advances on wearable computing could allow to integrate sensors in the cloth fibers, detecting signals directly from the shirt leaves. Some investigations for developing such prototype are already undergoing but the gesture set might need to be modified, in order to adopt only electrodes in the forearm. On the software side, the system is completed by a gesture classifier running on a PC, which processes the data received from the Bluetooth wearable device, extracting features and recognizing gestures through four Linear Discriminant Analysis classifiers (one for each gesture). A video of the system can be seen at the following address: https://youtu.be/YDL4QBbIZwY.

7.2.2.5 Evaluation

Since this prototype has been developed only for a feasibility study, mostly from a technological perspective, the aim of the evaluation was only assessing the system accuracy in classifying the selected gestures. Because of the complexity of the system setup, a usability evaluation has not been conducted. Eight people, aged within 23 and 31 years performed a total of 80 gestures in two session of 3 minutes. Participants were instructed through a display on the gesture to perform. This visual stimulus was used also to segment gestures, since no gesture segmentation was implemented in the system.

Figure 28. An evolution of the wearable system presented in this Section, which implemented online recognition of the four gestures. In the image, the system is recognizing the dragging upward gesture.

The collected gestures have been used to assess the classification accuracy of the system through a k-fold cross-validation (k = 10). The classifier obtained an overall accuracy of 94.55% (SD=3.77), with similar accuracies for the four gestures (min=92.98%, max= 95.51%). The first results obtained with this prototype were encouraging in terms or recognition accuracy and promoted further developments of the system, which are presented in Section 7.2.4.

7.2.3 WheelSense V2 (Embedded and Embodied)[20]

7.2.3.1 Explored TGIF Aspects

As for other WheelSense systems, also this system allowed to explore the use of tangible gestures for the *control* application domain, specifically, for controlling the infotainment system of the car. The interaction was designed using the *expert-driven* method, with a *reification* of new gestures performed on a given object, the steering wheel. Following field requirements, all gestures were designed including the *hold* component. Gestures were designed exploiting *embodied metaphors* and the *form* affordances provided by this object and chosen according to the recognition performances allowed by the selected *embodied and embedded* technology for recognizing gestures (pressure sensors). The technological approach chosen for the implementation of this system is the main difference with the previous system. Software-side, *machine learning* has been used to recognize gestures. The design space of TGIF explored through this system is represented in Figure A3 in Appendix A.

[20] This work has been published in [7] in 2013, in collaboration with Francesco Carrino, Stefano Carrino and Maurizio Caon.

7.2.3.2 Design Rationale and Methods

As for WheelSense V1, the aim of WheelSense V2 was to provide a richer user experience and a safer interface for the in-vehicle infotainment system. Following literature analysis and according to field requirement, the system should comply with the "eyes on the road, hands on the steering wheel" paradigm introduced by González et al. [79]. To this purpose, gestures should oblige the user to keep both hands on the steering wheel in the position suggested by the Swiss manual for driving school teachers [103]. This system has been designed following an *expert-driven* method, with a particular attention to the requirement of firmly holding the steering wheel with two hands. The system design was oriented towards an *embedded and embodied* approach, with sensors integrated in the steering wheel, to recognize gestures performed on its surface, exploiting the *form* affordances of the steering wheel to guide tangible gesture execution. The tangible gesture design was partially influenced by the approach chosen to develop the system. Since pressure sensors have been chosen to recognize gestures performed on the steering wheel, following the embedded and embodied approach, the tangible gesture choice was limited to only those gestures that imply the application of a pressure on the surface of the steering wheel.

The implementation of the system required several iterations for the optimization of the physical design and the sensor integration. In particular, the sensor position was crucial for optimizing the system performances according to the chosen gestures. Therefore, different sensor positions have been tested and compared in order to obtain a good quality of sensor data. At the same time, the ergonomics of the steering wheel should not be affected by the integration of the sensors. These aspects will be further discussed in the next section, as part of the interaction design.

7.2.3.3 Interaction Design: Tangible Gesture Syntax and Semantics, Feedback

Four tangible gestures have been designed for this system (Figure 29). In all the four gesture-object pairs, the object is obviously the steering wheel. A *hold* component is always present in the gesture syntax to comply with the requirement of having constantly both hands on the steering wheel while interacting with the system. Three gestures that have been designed for the wearable system, i.e., the fist squeeze for stopping music and the dragging up and down for browsing songs, have been adopted also for this system. For the fourth gesture, an index tap while holding the steering wheel (which introduces an additional touch component) replaced the index abduction gesture (move component) for starting music. Indeed, not only the index tap has a better referential association to the starting music command, but it also exploits better the affordances provided by the steering wheel. The sound generated tapping the wheel provides inherent feedback to the user and foster the association between the steering wheel and a musical instrument, which in this case becomes the interface with the infotainment system of the car. To this purpose, it is worth recalling that the purpose of this system was limited to the interaction with the music player, with the aforementioned commands for starting and stopping music and for browsing songs.

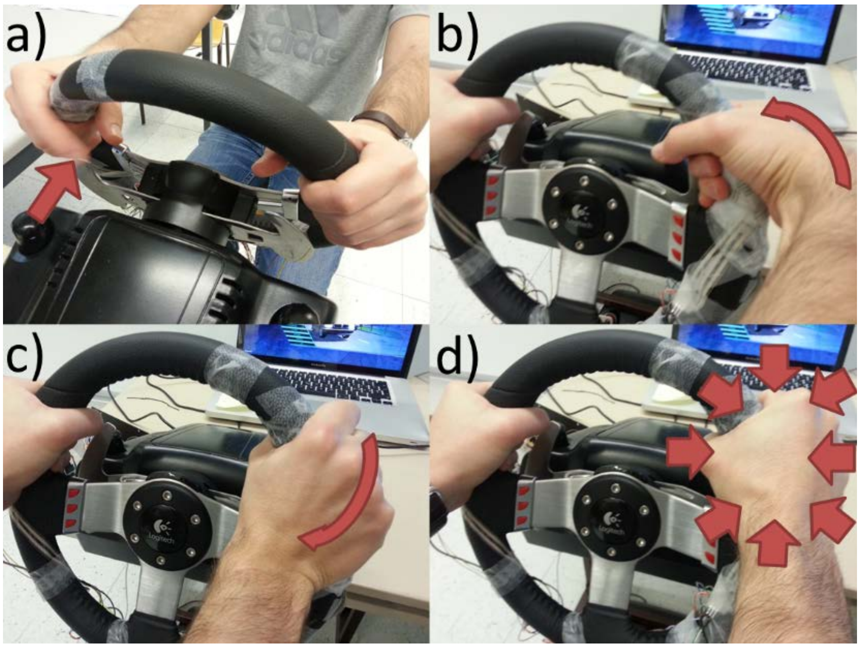

Figure 29. Representation of the four gestures: a) tap, b) dragging upward, c) dragging downward, d) squeeze.

Since tangible gesture execution relies on proprioception and on the inherent haptic feedback provided by the steering wheel, the visual demand for the driver to interact with the IVIS through this system is reduced, especially if coupled with proper feedback. Haptic feedback was excluded because, as stressed from Bach et al. [17], there is a risk of increased distraction if the secondary task competes on the perceptual resources required by the primary task. Indeed, haptic feedback coming from the road and perceived through the steering wheel has an important role in the driving task. As suggested by Wickens and Hollands [225], the perceptual resources needed for the secondary task can be distributed over other senses. As a result, to comply with the "eyes on the road" rationale, the only auditory feedback coming inherently from the IVIS was considered for this system, which is largely sufficient for the simple control tasks for which the system has been designed. Obviously, the purpose of the application, i.e. listening to music, ensures that the auditory channel is not disturbed, thus the feedback is effective.

As part of the interaction design, the object design was also important to ensure the best user experience in gesture execution. Indeed, the integration of the system in the steering wheel (a Logitech G27) required a particular attention for the object crafting. As discussed in Section 7.2.3.2, the sensor position was important to optimize recognition performance. At the same time, the surface of the steering wheel should remain smooth and soft, to allow proper operation of the steering wheel while driving, as well as the slick execution of tangible gestures on its surface. For the same reason, I avoided adding tactile cues on the steering wheel to help the user spot the right position for performing gestures, relying only on the original physical form and ergonomics of the steering wheel chosen for the integration. Transparent adhesive tape has been used to stick sensors on the steering wheel, which ensured smoothness of the surface and an easy rearrangement of sensors for further testing. Unfortunately, other design rationales, such as the visual appealing of the prototype, were sacrificed, resulting in a raw solution with wires hanging from the bottom of the steering wheel (see Figure 30). Such poor aesthetics can affect the perceived usability of the system, as shown by Sonderegger and Sauer [191].

7.2.3.4 Implementation

The sensing system implemented in this prototype is based on five Tekscan FlexiForce sensors with a range of 0-1 lb[21]. They are connected to an Arduino Duemilanove board that converts signals to the digital domain and sends measured data to a PC for further elaboration through a wired serial connection. Data are acquired with a rate of 50 Hz. Four sensors have been placed on the right side of the Logitech G27 Racing Wheel, for the right hand, and one sensor has been placed on the left side for the left hand. The sensor placement on the steering wheel is depicted in Figure 30.

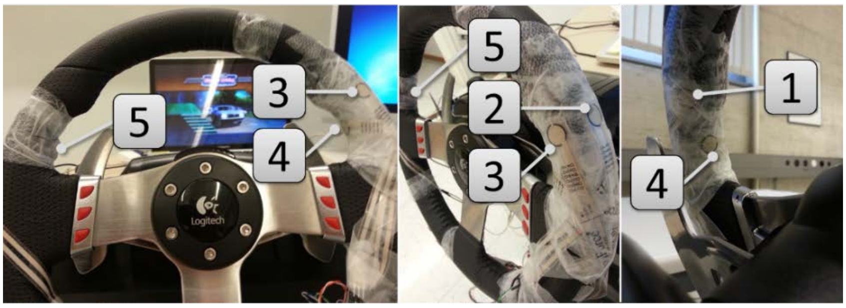

Figure 30. Placement of the 5 FlexiForce sensors

Sensor 1 is placed on the back of the steering wheel to recognize the tap gesture with the index finger. In a relaxed position, the hand generally covers the three other sensors. The wrist flexion and the wrist extension performed for the dragging upward and downward gestures uncover respectively Sensor 3 and Sensor 4. In order to segment gestures and avoid false positives that could occur while manipulating the steering wheel, an explicit segmentation strategy was adopted: the driver squeezes the left hand while gestures are performed with the right hand. This additional gesture for explicit segmentation is recognized through Sensor 5.

[21] Tekscan FlexiForce: http://www.tekscan.com/flexible-force-sensors

The raw data of the five sensors are elaborated in the PC using the ARAMIS Framework [44]. First, gestures of the right hand are segmented setting a threshold on the left hand sensor. Afterwards, the segmented data are used as input for a Hidden Markov Model (HMM) classifier. The HMM classifier was configured with 4 hidden states with forward topology and implemented the Baum-Welch algorithm to find the unknown parameters. The data supplied to the HMM classifier are modeled as time series (as depicted in Figure 31). The whole architecture of the WheelSense system is reported in Figure 32.

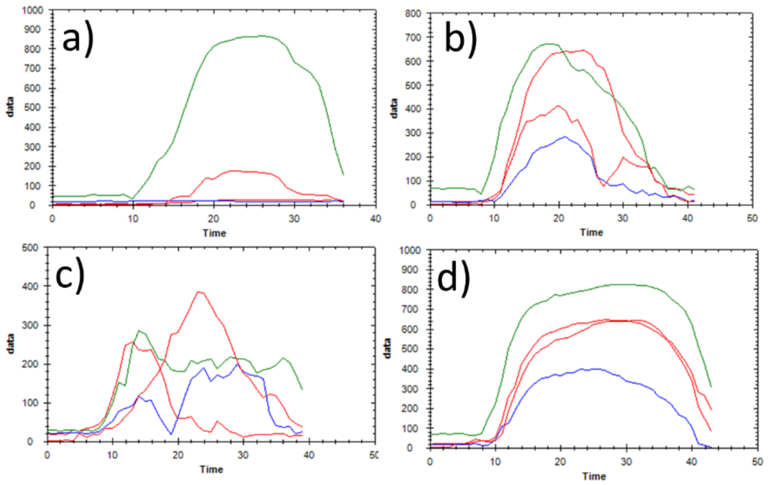

Figure 31. Representation of the temporal signals associated to the four gestures: a) is index tap, b) is dragging upward, c) is dragging downward and d) is squeeze.

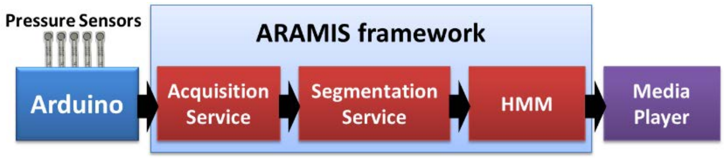

Figure 32. Block diagram of the WheelSense system architecture

A video of the system can be found at the following address: https://youtu.be/OfpAkJ3cUrE.

7.2.3.5 Evaluation

The aim of the evaluation was to assess the system accuracy as well as the usability of the system. The evaluation was composed of three phases: the evaluation of the gesture recognition accuracy in controlled conditions, the evaluation of the gesture recognition accuracy while the user was driving using a simulator and the evaluation of the system usability through a standard questionnaire. Eight users (six males and two females, aged 25 - 31) participated to these evaluations. The setup, depicted in Figure 33, is composed by a laptop that executes the recognition task and the City Car Driving

Simulator version 1.2[22]. The monitor on the right shows the results of the classification and it was used by the experiment supervisor.

Figure 33. A user participating at the evaluation.

During the first part of the evaluation process, the users were asked to perform each gesture 40 times, for a total of 160 gestures. The order of the gesture to be performed was chosen randomly and the user was guided by a graphical interface. The user was requested to rest at half of the recording phase. A 10-fold cross-validation test on the recorded data was performed. The resulting average accuracy was 87% and the standard deviation was 17%. The HMM classifier has been trained using the data recorded during the first phase. The 8 participants were asked to drive using the City Car Driving simulator and to interact with the IVIS through the gesture interface. In this case, the gestures were used to control a music player with the gesture semantics explained in the previous subsection. The experimenter requested the command that the participants had to perform; they had to remain focused on the driving task and to perform the gesture only when they were feeling confident, i.e., when the participants considered interacting with the IVIS as not dangerous. The total number of gestures that each user had to perform during the driving simulation was 40 (10 per type of gesture). The average accuracy was 82% and the standard deviation among users was 16%. In fact, during the experience, high variability between the different participants was noticed.

After the second phase, participants were asked to fill a System Usability Scale questionnaire (SUS) [31]. From the questionnaire the overall usability, perceived usability and the learnability had been calculated. The overall usability (calculated following the standard procedure) scored 84 points

[22] City Car Driving - Car Driving Simulator http://citycardriving.com/

out of 100 (standard deviation: 13); the perceived usability scored 82 points out of 100 (standard deviation: 12); the learnability scored 91 points out of 100 (standard deviation: 17). The last two factors were calculated as suggested by Lewis and Sauro in [127].

The two performance evaluations showed a high variability among users, which affected also the results of the usability evaluation. This high variability could be explained with the different hands position of the users during the interaction. In some cases, the right hand was not always positioned over the pressure sensors, which decreased consistently the quality and the strength of the acquired signals. Variations could also occur over time: for example, in one case, the system confused several times a squeeze with a dragging up gesture, because the participant was not pressing anymore on one of the sensors. This suggests that a robust system should be difficult to achieve without taking into account the changes in the users' behavior over time. An adaptive learning approach could be implemented in order to avoid this issue.

7.2.4 WheelSense V3 (Hybrid)[23]

7.2.4.1 Explored TGIF Aspects

As for other WheelSense systems, also this system allowed to explore the use of tangible gestures for the *control* application domain, specifically, for controlling the infotainment system of the car. The interaction was designed using the *expert-driven* method, with a *reification* of new gestures performed on a given object, the steering wheel. Following field requirements, all gestures were designed including the *hold* component. Gestures were designed exploiting *embodied metaphors* and the *form* affordances provided by this object and were chosen to obtain a common set of gesture that could be recognized through both the *wearable* and the *embedded and embodied* approaches. Indeed, the main purpose of this system was to explore the *hybrid* technological approach. Software-side, *machine learning* has been used to recognize gestures. The design space of TGIF explored through this system is represented in Figure A4 in Appendix A.

7.2.4.2 Design Rationale and Methods

The aim of this project was combining the two previous systems (wearable and embedded) presented in Section 7.2.2 and 7.2.3. This system has been designed according to the opportunistic synergy paradigm presented in Section 6.1.4. According to this paradigm, the two sub-systems should be able to work alone and when both are present in the final system, we should expect performances (in terms

[23] This section of the framework has been pushed in 2013 in [10] in collaboration with Stefano Carrino, Francesco Carrino and Maurizio Caon. It is worth mentioning that in this thesis I distinguish between the embedded and embodied approach and the environmental approach, while in the article [10] sensors embedded in the objects are considered as environmental sensors.

of recognition accuracy) that are equal or better than what can be achieved with the two sub-systems working alone. Moreover, the opportune combination of the two systems should bring interaction advantages also from a qualitative point of view, which should be evaluated according to the design parameters presented in Section 6.1.4. In order to obtain such a hybrid system, the two previous systems have been slightly adapted. Indeed, in order to be able to work one independently from the other, the two sub-systems should be able to recognize the same gesture set and both should allow proper gesture segmentation. The gesture set design followed the rationale of the previous systems, with the purpose to comply with the paradigm “eyes on the road, hands on the steering wheel” to minimize the visual demand (thanks to inherent haptic feedback provided by the steering wheel), and the cognitive load for the execution of gestures, (thanks to the embodied metaphors used for designing the semantic relationships between gestures and commands).

Moreover, the purpose of this study was to compare different strategies for the segmentation of gestures and for the classifier fusion. In particular, explicit segmentation (performed by the user with an additional gesture) and implicit segmentation (automatically performed by the system) have been compared. Since implicit segmentation can lower the cognitive load of the user and can also free motor resources, evaluating the performance of a system based on implicit segmentation is crucial to understand if gestures can be automatically segmented by the machine without affecting the recognition accuracy of the system and, in the end, its usability.

7.2.4.3 Interaction Design: Tangible Gesture Syntax and Semantics

The gesture set includes the four tangible gestures designed for the embedded and embodies system (index tap to start music, dragging upward/downward to browse songs, hand squeeze, to close the player), plus an additional gesture, a hand press (push) in the frontal part of the steering wheel to pause music). The five gestures are shown in Figure 34. The interaction design principles are similar to those discussed for the previous two systems. Feedback was not present in this system since the user evaluation was focused only on recognition accuracy.

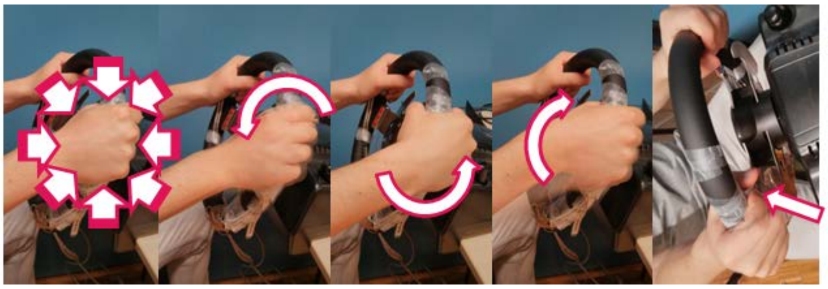

Figure 34. The five gestures for the hybrid WheelSense system. From left to right, hand squeeze, dragging downward, dragging upward, hand press, index tap.

7.2.4.4 Implementation

The system has been implemented following the hybrid approach, integrating the two subsystems presented in Section 7.2.2 and 7.2.3, thanks to the ARAMIS framework [44]. The wearable system has been slightly modified: the electrodes for EMG sensing have been arranged differently on the right arm in order to detect the muscular activity of the *Extensor Carpi Ulnaris*, *Flexor Carpi Ulnaris*, *Extensor Digitorum*, *Flexor Carpi Radialis* and *Palmaris Longus*. One electrode was stick also in the left arm to detect the hand squeeze gesture (for explicit gesture segmentation). The embedded and embodied system was left almost unchanged, with only a slight modification of Sensor 3 position (see Figure 30) to ensure optimal detection of the hand press gesture. Two different architectures of the hybrid system were developed: one based on explicit segmentation (executed manually by the user with the left-hand squeeze) and the second one based on implicit segmentation (automatically determined by the system). The architecture of the hybrid system with explicit segmentation is shown in Figure 35.

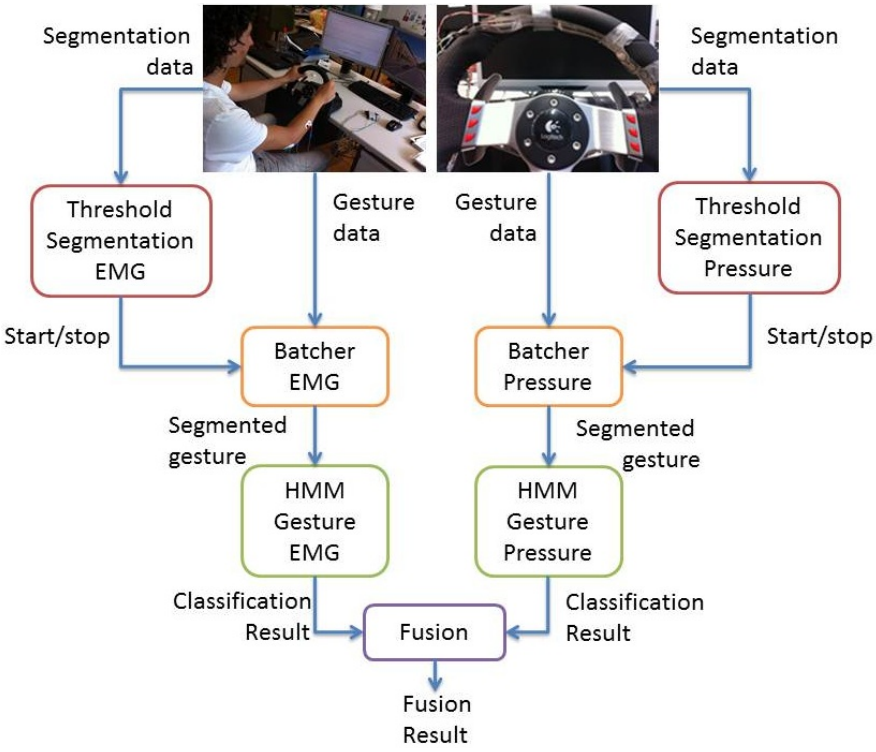

Figure 35. Architecture of the hybrid system with explicit segmentation

In this architecture, each sub-system provides different types of information flow: segmentation data and gesture data. Information from EMG and pressure sensors is analyzed in parallel. Each sub-system segments the gesture data coming the sensors for the right hand according to the segmentation data received from the sensor for the left hand. Indeed, while the user is squeezing the steering wheel with the left hand, a gesture performed on the steering wheel by the right hand is meant as command and must be recognized.

The gesture data is segmented using thresholds that are fixed in a user-dependent calibration phase. In particular, the segmentation module adopts a hysteresis-based approach, in which the activation and deactivation thresholds are calculated taking into account the respective standard deviations of the signal values measured when the user is driving and when is performing the *Squeeze* gesture (see Figure 36). Therefore, the segmentation is enabled when the signal (the root mean square for the EMG and the raw data for the pressure sensor) exceeds the *activation threshold* and is disabled when the signal drops below the *deactivation threshold*.

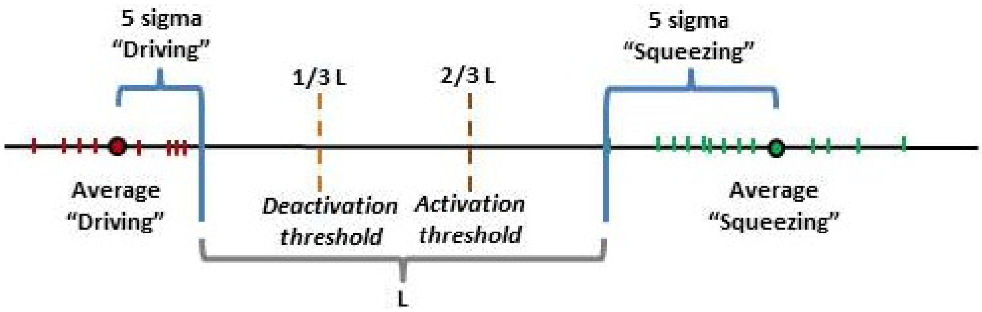

Figure 36. Hysteresis approach for the calculation of segmentation thresholds

The Batcher modules start to accumulate the data on the start signal. With the stop signal, the collected information is sent to the classification algorithm. The Hidden Markov Model (HMM) classifiers were configured with 4 hidden states with forward topology and implementing the Baum-Welch algorithm to find the unknown parameters. Finally, the outputs of the HMMs are merged in the fusion module. The fusion module performs the fusion of the segmentation results of the two sub-system according to the three different strategies presented in Section 6.3.3, as well as the fusion of classifiers results according to the 10 strategies presented in the same Section.

The explicit segmentation uses the same signals (gesture data) for gesture recognition and segmentation. Practically, it involves a reduced cognitive load on the user who has to perform only right hand gestures on the steering wheel. The architecture, presented in Figure 37, is slightly different from the one presented in Figure 36.

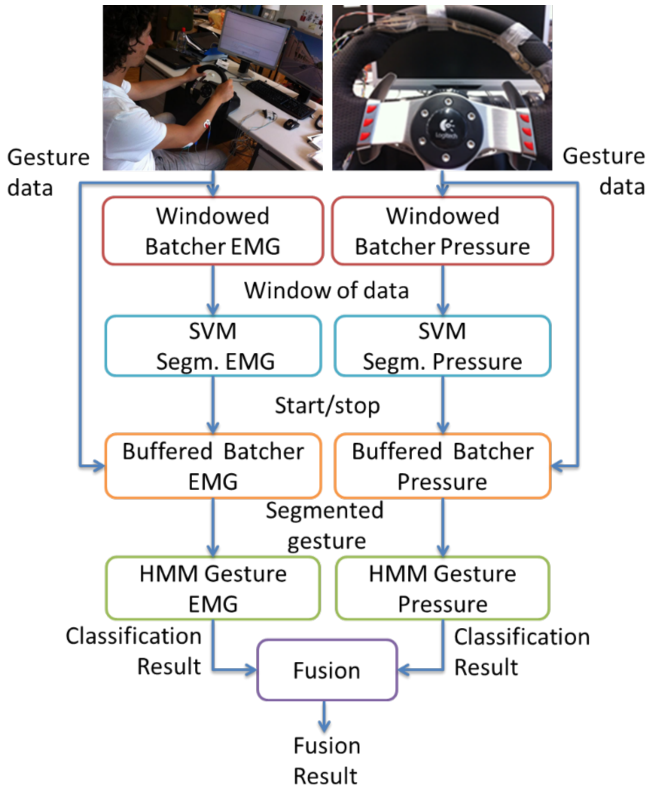

Figure 37. Architecture for the hybrid system with implicit gesture segmentation

In order to spot the start and the end of a gesture the system analyzes the data flows using a windowing approach. The window size is 3 samples with an overlap of 33% for the EMG and 10 samples with an overlap of 50% for the pressure sensors. These values have been selected empirically during a pretest phase. For each window, a SVM classifier estimates if the user is performing a gesture or is simply driving. The SVM algorithm uses a Gaussian kernel in which an appropriate value of sigma is obtained using the approach explained in [40] as implemented in the Accord.Net framework [44]. Since the segmentation events are fired after a window of data, there is the risk of losing important information. The buffered batcher modules allow queuing data and avoiding this loss. The following part of the schema is the same as in the explicit segmentation approach, with HMMs classifiers analyzing the data for gesture recognition.

7.2.4.5 Evaluation

The aim of this study was to compare the gesture recognition accuracy of the different systems (wearable, embedded and embodied, hybrid) of the two segmentation strategies (implicit and explicit), and of the different fusion strategies. Moreover, it aims at discussing the interaction advantages and drawbacks of combining two different systems according to design parameters discussed in Section 6.1.4.

9 users (1 female), aged 24-31, tested the system according to the following protocol. The tests were conducted in a controlled laboratory setup, with a Logitech G27 steering wheel in front of a display that informed the user on the primary tasks (steer the wheel, accelerate, brake, etc.) or secondary tasks (i.e., the 5 tangible gestures) to perform. First, the participants had the possibility to familiarize with the system, for about five minutes. Then, each user had to follow a calibration phase of the system. This phase aims at determining the thresholds for the explicit segmentation, measuring the average strength applied by the user on the steering wheel and the average electrical activity of the muscles, while driving and while performing the squeeze gesture. Such signals are user-dependent and vary according to the user driving habits and muscular development (with significant differences between genders).

For the next step, each user was asked to perform 30 gestures for each of the five gesture classes (hand squeeze, dragging upward, dragging downward, index tap, and hand press) in a random order. Between the execution of the different gestures, the users were asked also to perform normal driving actions on the steering wheel: turn left, right, accelerate, brake or stay in a rest position. The acquisition protocol consists of the simultaneous collection of data for the implicit and explicit segmentation systems. The users were asked to squeeze the steering wheel with the left hand to perform the explicit segmentation. This information was then used only by the manual segmentation modules. The whole acquisition process was performed in two sessions of about 15 minutes in order to allow the users to rest. 70% of the collected gestures were used as training and cross-validation sets; the 30% of the gestures were used as test set. I used a k-fold (k=10) cross-validation on the training in order to calculate the confusion matrix for each classifier and to select the weights for the fusion module (cf. Section 6.3.3).

To compare the performances of the different systems, segmentation strategies and fusion strategies, the F1-score and the recognition accuracy in the classification have been calculated. Table 3 presents these values for the different classifier fusion methods and segmentation strategies. Pressure and EMG rows present the results of the single classifiers.

Table 3. Fusion results (μ and σ) F1 score and accuracy. In red the best single classifier; in bold the best fusion methods for each segmentation strategy. Please refer to [10] for additional details on the different fusion strategies.

(μ|σ)	Explicit Segment. OR		Implicit Segment. OR		Explicit Segment. AND		Implicit Segment. AND		Explicit Segment. ADAPTIVE		Implicit Segment. ADAPTIVE	
Method	**F1**	**A**	**F1**	**A**	**F1**	**A**	**F1**	**A**	**F1**	**A**	**F1**	**A**
Pressure	0.72|0.16	0.79|0.14	0.66|0.20	0.76|0.17	0.72|0.17	0.76|0.15	0.65|0.15	0.76|0.13	0.72|0.16	0.76|0.13	0.65|0.16	0.76|0.12
EMG	0.85|0.12	0.86|0.13	0.73|0.19	0.75|0.19	0.85|0.12	0.85|0.13	0.75|0.19	0.74|0.19	0.85|0.11	0.85|0.12	0.74|0.19	0.72|0.19
SR	0.79|0.13	0.85|0.11	0.70|0.15	0.84|0.09	0.77|0.18	0.72|0.24	**0.78|0.12**	**0.73|0.14**	0.81|0.12	0.82|0.12	**0.74|0.11**	**0.77|0.11**
NB	0.79|0.14	0.84|0.10	0.71|0.15	0.84|0.08	0.74|0.21	0.70|0.26	0.76|0.12	0.70|0.14	0.78|0.11	0.79|0.11	0.71|0.16	0.74|0.15
NBW	0.80|0.14	0.85|0.10	0.71|0.15	0.84|0.08	0.75|0.21	0.71|0.27	0.77|0.13	0.71|0.14	0.80|0.12	0.80|0.12	0.71|0.15	0.74|0.15
MCC	0.81|0.14	0.87|0.09	**0.72|0.14**	**0.86|0.09**	0.80|0.20	0.75|0.26	**0.78|0.13**	**0.73|0.15**	0.84|0.11	0.85|0.10	**0.74|0.13**	**0.77|0.12**
MCC+	0.81|0.15	0.87|0.10	0.69|0.16	0.83|0.14	0.78|0.20	0.73|0.25	**0.78|0.16**	0.72|0.17	0.83|0.12	0.85|0.11	0.73|0.16	0.76|0.15
SR*	0.79|0.13	0.85|0.11	0.70|0.15	0.84|0.09	0.77|0.18	0.72|0.24	**0.78|0.12**	**0.73|0.14**	0.81|0.12	0.82|0.12	**0.74|0.11**	**0.77|0.11**
NB*	0.79|0.14	0.84|0.10	0.71|0.15	0.84|0.08	0.74|0.21	0.70|0.26	0.76|0.12	0.70|0.14	0.78|0.11	0.79|0.11	0.71|0.16	0.74|0.15
NBW*	0.80|0.14	0.85|0.10	0.71|0.15	0.85|0.08	0.75|0.21	0.71|0.27	0.77|0.13	0.71|0.14	0.80|0.12	0.80|0.12	0.71|0.15	0.74|0.15
MCC*	**0.83|0.15**	**0.89|0.10**	**0.72|0.14**	**0.86|0.09**	0.80|0.21	0.75|0.26	**0.78|0.15**	**0.73|0.16**	**0.85|0.13**	**0.86|0.12**	**0.74|0.14**	**0.77|0.13**
MCC+*	0.82|0.15	0.88|0.10	0.71|0.18	0.84|0.15	**0.81|0.21**	**0.76|0.27**	0.77|0.17	0.72|0.18	**0.85|0.13**	**0.86|0.12**	0.72|0.16	0.76|0.15

From Table 3, it can be observed that in general the wearable system performed better than the embedded and embodied system. The best results have been achieved with explicit gesture segmentation, the ADAPTIVE segmentation fusion strategy and the MCC* classifier fusion method. Such configuration performs better of the other fusion methods and, in particular, equal or better than the best standalone classifier.

Finally, the manual segmentation performs in average about 6% better than the automatic segmentation. In order to understand which segmentation approach should be chosen in an interaction system, the trade-off between the cognitive load on the user and the effect of a lower accuracy of the gesture recognition should be evaluated.

The quantitative evaluation conducted on this particular dataset shows good performances in terms of accuracy and F1-score. However, this study allowed investigating also more qualitative features of a synergistic paradigm in the context of the interaction in a car, i.e., the design parameters presented in Section 6.1.4. A direct consequence of the wearable paradigm is that the *interaction area* is no more limited to some spots in the car but can be extended to the whole car. Once the communication between the wearable and the embedded and embodied systems is established, it is possible to control the IVIS everywhere in the car. Gestures can be user dependent. *Personalized* configurations and profiles can be shared from the wearable system to the embedded and embodied one. In addition, a wearable system can be designed to share the user information with an unknown system. For example, in a car-sharing scenario, the information of the user stored in the wearable system can be used to configure a new car automatically, binding the personalized gesture vocabulary to the system controls. With a synergistic system, an easy way to increase *consistency* is to adapt online the weights used in the fusion module to achieve better performances. For example, it is possible to penalize or award a classifier directly decreasing or increasing the weights used by the SR or the MCC methods. The car is an interaction milieu that is generally considered as private. Therefore, in this specific context the needs of *private interaction* and intimate interfaces are reduced, even though the wearable components represent a good solution for this problematic. The embedded and embodied component of a synergistic system can provide *localized information* to the driver taking into account the specificities of the car. For example, the windshield can be used to display information to the user that does not need to move the gaze from the road conserving the safety of the driving. As mentioned before, the 5 pressure sensors were placed in specific regions of the external ring of the steering wheel to have *localized controls* that should help the driver to keep the hands on the steering wheel, at least while interacting with the IVIS. The embedded and embodied sensing can be directly integrated in the car exploiting the existing processing and power resources, which, in this case, can be much higher than those of the wearable system. A synergistic paradigm allows the wearable system to take advantage of the vehicle *resources availability*. In fact, even if the wearable components still require energy to sense the information, the processing can be deployed at vehicle side. The downside of this approach is that the transmission of the information can be highly demanding in term of energy and may slow down the whole processing system. Therefore, accurate analyses should be performed case-by-case. The presence of different gesture lexicons for the wearable and the embedded and embodied paradigm can be treated as *resource management*. Commands linked to secondary tasks can be integrated as wearable components and made available for the driver as well as the passengers of the vehicle; on the other hand, primary commands, that can affect passengers' safety, can be made accessible only to the driver by localizing the control in a particular spot of the vehicle.

The previous examples can be used as guideline to tune the design parameters of an in-vehicle opportunistic interaction system in order to take advantage of the best features of both the embedded and embodied and the wearable systems.

7.2.5 Gesture Elicitation on the Steering wheel[24]

7.2.5.1 Explored TGIF Aspects

This study has been conducted to explore the use of tangible gestures for the *control* application domain, specifically, for controlling the infotainment system of the car, through a gesture elicitation. The results of this study has been used to design the system presented in Section 7.2.6 (WheelSense V4). Therefore, with this study, I investigated the shift from an *expert-driven* to a *user-driven* method. The gesture elicitation has been conducted to explore all the possible tangible gestures that can be performed in relation to the steering wheel object (a Logitech G27). In this case, the *hold* component was not a given requirement, but a *touch* contact was. Elicited gestures still exploited the *form* affordances provided by this object, but in this case, the steering wheel form was exploited more often as a spatial reference rather than for grasping affordances. In this section, I present the gesture elicitation design, the results and the consideration emerged from the analysis of the results.

7.2.5.2 Design rationale and methods

In the three previous systems, the choice of gestures followed the rationale of complying with the safety requirement to keep both hands on the steering wheel. The gesture sets of the previous systems were designed by experts, taking into account also the technology chosen to recognize gestures, in order to maximize the gesture recognition accuracy. As discussed in Section 5.2, an alternative approach for designing the gesture set is asking directly to the users which kind of gestures they would like to perform and for which commands. Indeed, gesture elicitations allow designing gestures that are more intuitive for the user, easier to guess and generally more appreciated than gestures designed by experts [144]. To this purpose, in order to design tangible gestures to interact with the infotainment system that are easier to learn and remember, a study for eliciting tangible gestures on the steering wheel has been conducted.

The study was conducted among younger drivers, in order to understand better which is the impact of the previous knowledge of modern device interfaces on the elicited gestures (cf. discussion on legacy bias presented in Section 5.2). 40 participants, 34 females and 6 males, aged 19-38 (average 22, SD: 3.3) participated to the study. 87.5% of our participants were right-handed. All the participants have a driving license and previous experience in interacting with touch interfaces such as smartphones or tablets. In a pre-elicitation phase, all the users participated to an experience (presented in Section 7.2.6) where they had to drive for about 30 minutes in a simulated environment. The setup was composed of a driving simulator running in a three-monitor configuration and a Logitech G27 steering

[24] This work has been pushed in 2014 in [9] in collaboration with Stefano Carrino, Francesco Carrino, Maurizio Caon and the department of Psychology of the University of Fribourg.

wheel (see Figure 47). During this phase, participants were asked to perform several interactions with a responsive In-Vehicle Information System (IVIS): handling the music (select songs and regulating the volume), traffic alerts, incoming calls, etc. Participants were separated in two groups. Group 1 interacted with an IVIS based on vocal interaction and a HUD, the commands used for the vocal interaction were the same of the gesture elicitation study (select, back, next, previous, volume up, volume down). Group 2 interacted with a touch dashboard with UI buttons. Both interfaces were based on the same menu-based structure. In both interfaces, the current item was in the center of the screen, next and previous items were respectively on the right and on the left of the current item and volume information was displayed just over the current item.

For the elicitation phase, participants were instructed through a video and were asked to elicit gestures in a given time (30 seconds for each gesture) while thinking aloud. Participants were recorded with two cameras and two people annotated the elicited gestures, according to the gesture definitions mentioned in Section 3.2. Participants were sitting in front of a Logitech G27 steering wheel; a laptop was used to show the instruction video. Since this study aimed at analyzing gestures on the steering wheel, some constraints to the interaction were fixed. Only gestures that implied contact with the steering wheel, including the external ring and the spokes, were allowed (either implying pressure, movements of the fingers or of the hands). On the contrary, the usage of paddles and buttons as well as mid-air gestures was forbidden. In order to allow the participants to focus only on gestures, the elicitation study was not conducted while driving. In this manner, participants had more cognitive resources at their disposal for the gesture elicitation.

Group 1 was asked to elicit gestures for three given pairs of commands of a menu-based interface in a HUD. The command pairs were: select/back, next/previous, volume up/volume down. The same command set was used in the vocal interface during the pre-elicitation phase.

In contrast, the second group was asked to elicit 6 gestures with no predefined commands associated and the participants had to define the system behavior (the associated commands) for each gesture. This second approach was chosen to give more freedom to the users for the definition of gestures. By dividing participants in these two groups, I aimed at investigating whether specific assigned functions have an impact on the elicited gestures or not.

The study had multiple purposes:

- Individuating the locations on the steering wheel that are more suitable for performing gestures

- Understanding the type of gestures preferred by users, as well as with which parts of the body are usually performed.
- Evaluating the impact of legacy bias
- Evaluating the impact on creativity of eliciting gestures for predefined commands
- Evaluating how much participants were concerned by safety constraints

The recorded videos have been annotated according to the aforementioned criteria. In particular, for the legacy bias criterion, any reference to existing interfaces given by participants during the think-aloud protocol was annotated. Gesture types have been annotated according to the nomenclature proposed in Section 3.3. In particular, gesture types were distinguished between tap (only the touch component is present), swipe (unidirectional touch+move gesture), stroke (multidirectional touch+move gesture), press (gesture that include a pressure attribute) and squeeze (a press gesture performed with a closed fist).

7.2.5.3 Results: Gesture locations

The steering wheel has been divide in 44 zones and unfolded in Figure 38 to show in a planar view also the posterior and lateral zones.

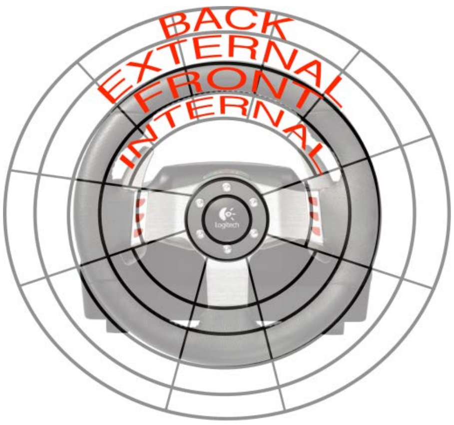

Figure 38. Steering wheel unfolded to show the four sides and divided into 44 zones

Zones have been defined through the visual and tactile cues that the Logitech G27 offered to the users. It is worth noting that a gesture may be performed on more than one zone: in this case, all the touched zones for that gesture have been counted for the generation of a heat-map. Figure 39 presents the heat-maps obtained for the gestures elicited by Group 1 and Group 2. The two heat-maps show that few zones of the steering wheel are left untouched, especially for Group 2, where only one zone was left untouched. Both left and right spokes have been used for more than 10 gestures for each group. The heat-map of Group 1 is more symmetric, probably because participants were asked for pairs of specular commands. Group 2 participants, who were free to define their own commands, performed more gestures on the right side than on the left side, probably because most of them were right-handed. The preferred side of the steering wheel for Group 2 was the front side with 86 occurrences (without spokes), followed by the external side (65 occurrences), while the Group 1 preferred the external side (74 occurrences) to the front side (44 occurrences, without spokes). It is worth noting that in order to reach some locations, users had to temporarily detach the hand from the steering wheel, which was quite natural for some participants, while others preferred to focus on zones next to the normal holding position for security concerns.

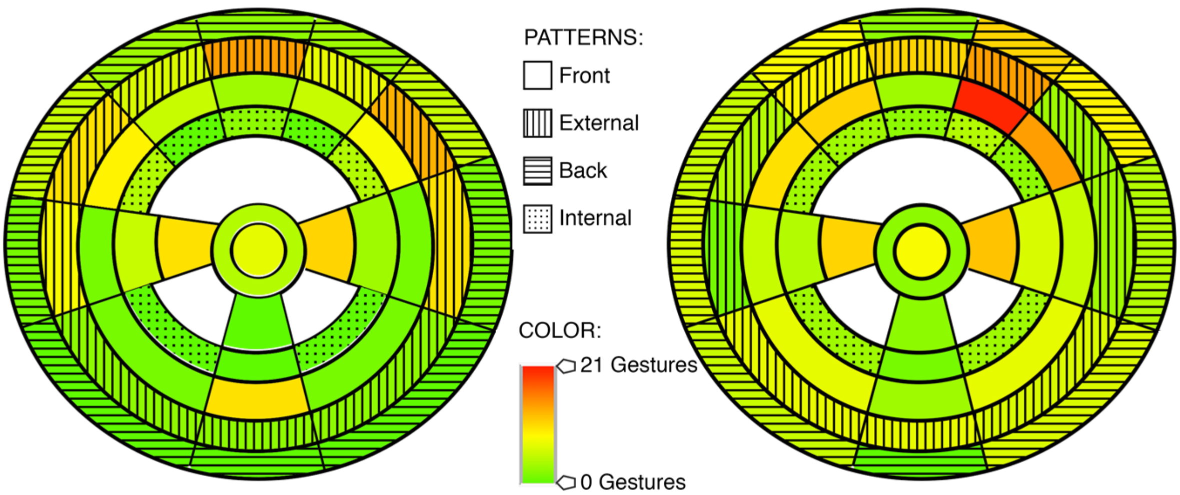

Figure 39. Heat-maps for gesture locations for Group 1 (left) and Group 2 (right).

Moreover, a frequent association between the gesture location and visual cues on the steering wheel has been evidenced from the analysis of the recorded videos. The spokes but also the center of the G27 Logitech steering wheel offered both tactile and visual cues that stimulated in few cases particular gestures, such circling around the central bezel. Moreover, since the G27 steering wheel has a small diameter (27cm), some participant tried to reach the center without leaving the hand from the correct holding position. An obvious implication of these observations is that an appropriate design of the steering wheel could be used to enhance safety for in-vehicle gesture interaction.

7.2.5.4 Results: Gesture syntax

This sub-section presents the analysis of body parts and gesture types chosen by the participants during the elicitation. There is no significant difference between the two groups for the body parts used to perform the gesture on the steering wheel. Figure 40 shows a clear preference in both groups for gestures performed with a single hand and the thumb. This result could suggest that the preferred part of the body for performing gestures does not depend on the associated function. However, between the two groups there is a statistically significant difference between gestures performed with “single hand” and “index” (considering a 95% confidence interval). Indeed, single hand gestures have been chosen by the first group 27±12% more often than the second group. Conversely, there is an opposite trend for gestures performed with the index finger. This probably means that gestures performed with the single hand were more suitable for the proposed command set (select/back, next/previous and volume up/volume down). Furthermore, it is worth noting that, in average, each participant tended to use a limited number of different body parts for all the gestures. In particular, participants in Group 1 used 1.65±0.13 body parts while members of Group 2 used 2.4±0.15 body parts. The difference in terms of different body parts adopted can be explained by a larger freedom in the choice of gestures for Group 2.

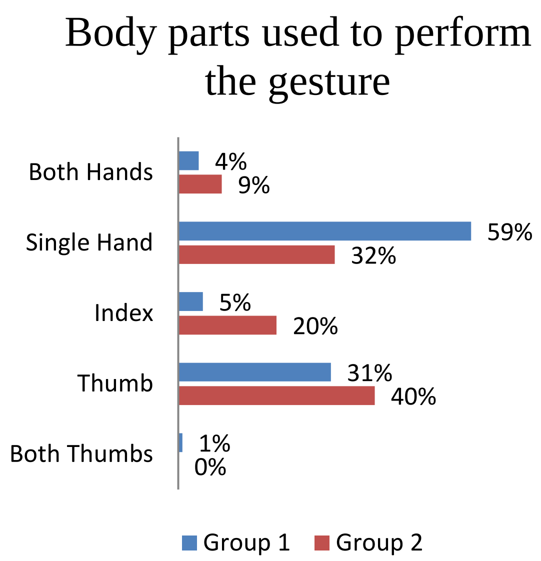

Figure 40. The bar chart summarizes the proportion of the body parts used by the participants to perform the various gestures.

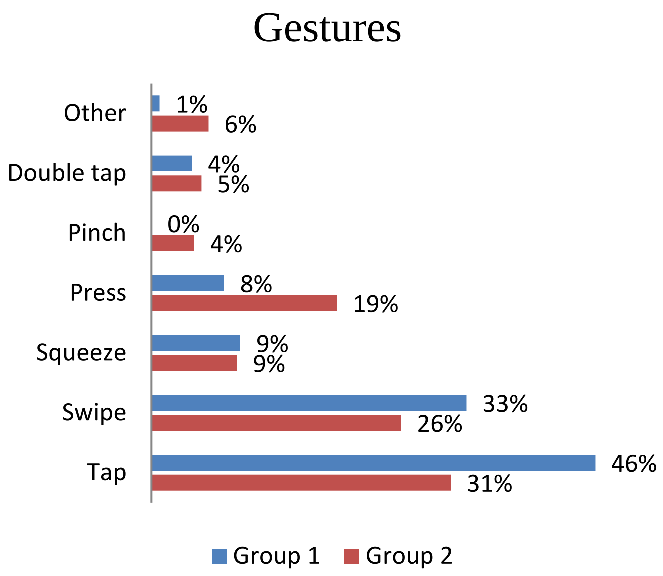

Figure 41. The bar chart summarizes the gesture adopted by the participants.

The analysis of gesture types (Figure 41) reveals that the two groups have a similar trend with an evident preference for the “tap” and “swipe” gestures (especially in Group 1). As already seen in the analysis of body parts, the gestures not related to a specific command set, performed by Group 2, show a larger variability among gestures, with more occurrences in the class “other”. This empirical observation is statistically confirmed: the members of Group 1 completed the task using in average 2.45±0.15 gestures while member of Group 2 (without a command set to take into account) used a more varied set of gestures resulting in 3.3±0.23 gestures proposed per participants. Furthermore, there is a statistically significant difference between the two groups for the adoption of the “tap” and the “press” gestures.

Concerning the type of gesture that participants would like to perform, an important insight evidenced by the think-aloud protocol is the users’ need for direct manipulation gestures. Indeed, while this study was conceived to elicit discrete gestures for specific commands that the system would execute only when the gesture is completed, several participants proposed the possibility to fine tune system functions such as the volume with a continuous gesture, instead of performing repeated “volume up/down” commands. This implies that some of the gestures that in this study has been annotated as swipes were actually performed as more controlled “drags”, which would require a higher sensor resolution in order to be accurately recognized by the system. Therefore, direct manipulation gestures for the car infotainment system should be also considered for a proper design of a tangible gesture interactive system in the car (cf. Section 5.6). An elicitation study for such gestures has been presented by Pfleging et al. using a touchscreen integrated in the steering wheel [160].

The results of the elicitation study for the gesture syntax can be summarized deriving a taxonomy for the most common tangible gestures elicited by both groups. According to the previous analysis, gestures have been classified according to the body parts used to perform the gesture and the gesture types. Figure 42 shows the most popular gestures and their occurrences, classified according to the taxonomy. Three macro categories can be identified: tapping, swiping and applying pressure. I excluded from the taxonomy less popular gestures, i.e., gestures with less than 4 occurrences. It is worth noting that gestures presented in Figure 42 can be executed in different zones of the steering wheel, either on the external ring or on the spokes. The location where these gestures can be effectuated is likely to depend on the type of command associated and/or how feedback is displayed. Moreover, some gestures of Figure 42 can be performed either while firmly holding the steering wheel or releasing the hand from it, depending if the gesture will be performed next to the driver's hands position or not (e.g., index tap, thumb tap or hand squeeze).

It is worth noting that according to the TGIF syntax, all tap gestures represented in Figure 42 include the touch component (this was indeed a requirement imposed to the participants to the elicitation), all swipe gestures include the move component and all the "pressure" gesture obviously imply a pressure attribute. As mentioned before the hold component can vary also across gestures of the same category, but it can be summarized that all hand taps, presses and swipes *do not* include a hold component, all squeeze gestures *do* include a hold component and thumb and index taps, presses and swipes *may* include a hold component (often they do).

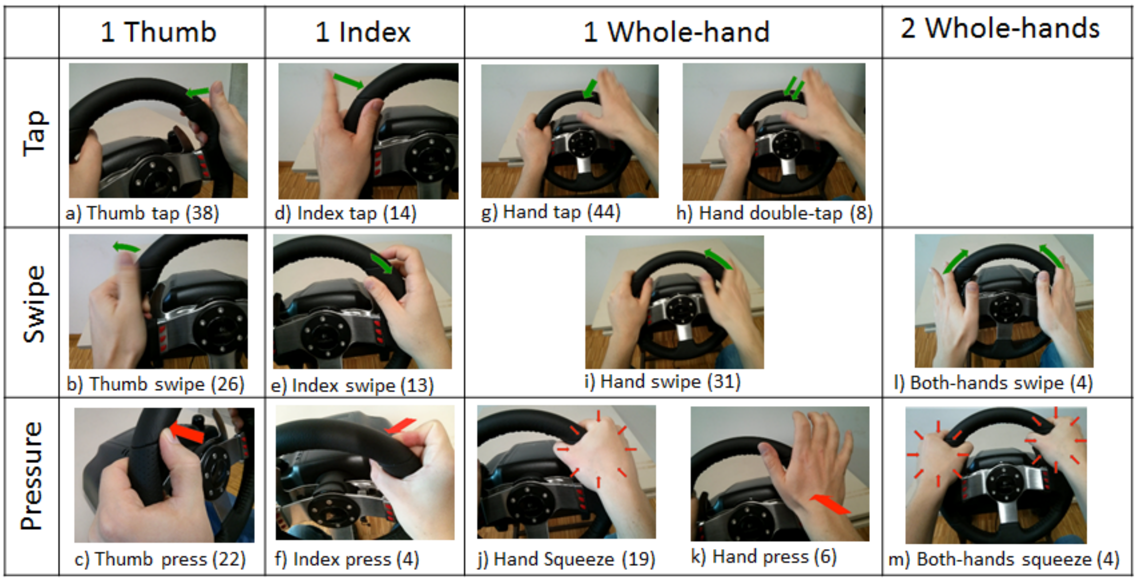

Figure 42. The taxonomy of gestures on the steering wheel derived from the elicitation study.

7.2.5.5 Results: Gesture semantics

In order to understand if the different assigned commands can influence the chosen body parts or gesture types, the gestures elicited by Group 1 were analyzed more in depth. Since the collected data are not enough to perform a chi-square analysis (the sample size condition is not verified), an inference via simulation approach for proportion was used, comparing the results of the simulation with the data actually collected to see if there were statistically significant differences (with a significance level, α=5%).

The first evident result is that "symmetric" commands implied symmetric gestures, i.e., the paired gestures select/back, next/previous and volume up/volume down resulted have almost identical (pairwise) results for the chosen body parts and gesture types (Figure 43).

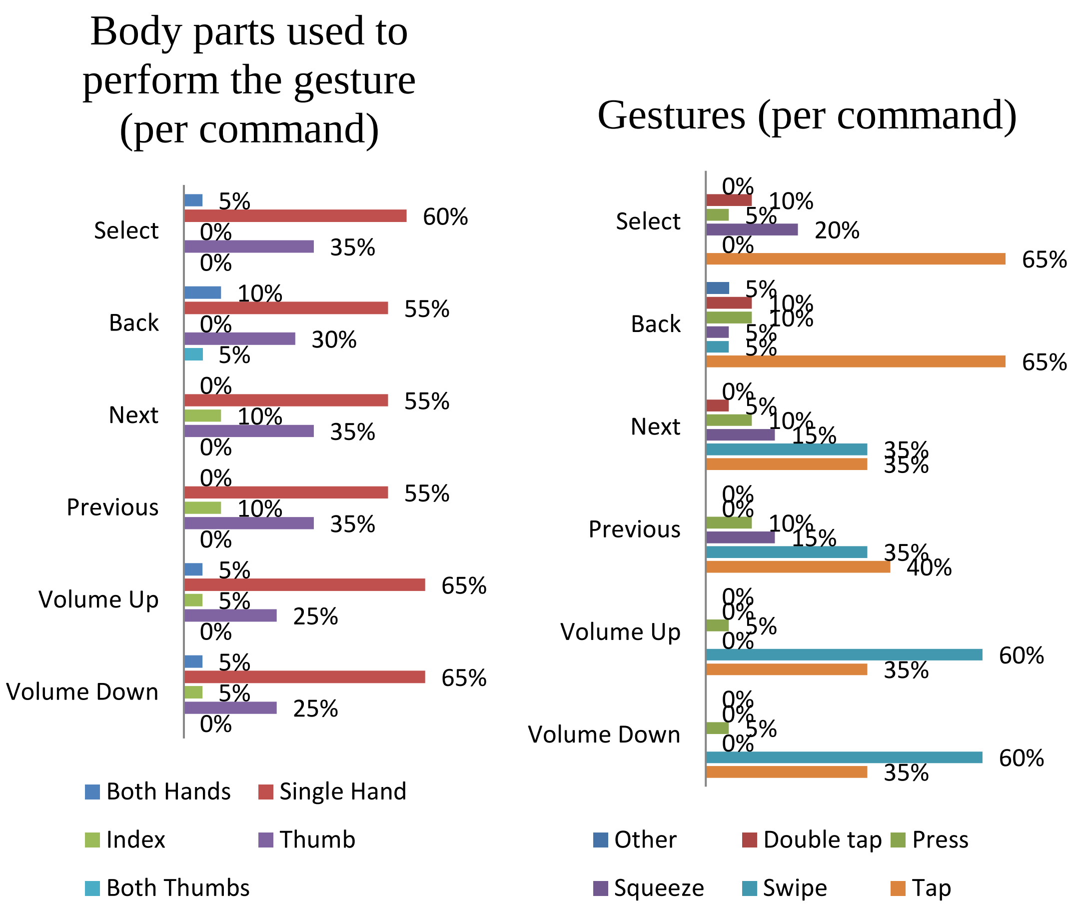

Figure 43. Analysis of gesture semantic associations for Group 1

The analysis of Figure 43 shows that the choice of the body part used to perform the gesture is almost unrelated to the associated command. The users showed a constant preference for gestures performed with a single hand, followed by gesture performed with the thumb. Other body parts (both hands, index, both thumbs) were less employed by the users.

On the contrary, the choice of the gesture type resulted strongly related to the associated command. The adoption of “swipe” significantly changed among the different commands: it was almost not used for the select/back commands, has been used often for next/previous commands, and much more appreciated for volume up/volume down commands (60% of these commands were performed as a “swipe” gesture). On the contrary, the “tap” gesture was appreciated for the select/back commands (65% of preferences), decreasing to around 35% for the other commands.

The function associated to a gesture is clearly linked in this dataset to a spatial reference in the gesture execution. Indeed, 100% of users have used the right hand, the right part of the steering wheel or gestures towards the right to perform the gesture associated to the “Next” command. Conversely, the “Previous” command is associated to movements to the left or gestures on the left side of the steering wheel. I measured very similar results with the “Volume Up” and the “Volume Down” gestures and the related parts of the steering wheel or direction of the gesture (“Volume Up”: 92.9% in the top part of the steering wheel or in the up direction. “Volume Down”: 95.2% in the bottom part of the steering wheel or in the down direction). While the spatial reference for next and previous commands can be easily linked to the behavior of the menu in the head-up display, “Volume Up” and “Volume Down” spatial references can find their root in the behavior of most common interfaces, as well as on the name of the command. Interestingly, comparing only the left and right components of the interaction, I noticed also that “Volume Up” gestures are performed for the 86.6% in the right part of the steering wheel and the 71.4% of the “Volume Down” gestures are in the left part of the steering wheel, showing a strong correlation between up and right, and down and left.

7.2.5.6 Results: Safety concerns, legacy bias

In order to understand how much users were concerned by safety, the number of gestures proposed by the participants that were performed holding firmly the steering wheel against the gestures that required the participant to leave one hand from the correct driving position has been calculated. Only 41% of the gestures proposed by Group 1 were performed in the holding position, against the 54% of Group 2. This is coherent with the previous results; in fact, in the first group there is a higher use of gestures performed with the whole hand, especially hand taps and swipes, which typically required shifting or temporarily removing the hand from the correct driving position.

Observing people during gesture elicitation, different behaviors has been found among users. Some people were concerned about safety and were looking for gestures to perform without releasing the hand from the steering wheel. Other people, instead, were looking for very simple gestures and were not much concerned by this safety risk. In few cases, gestures even required leaving both hands from the steering wheel, such as, both-hand swipes, or both-hand squeezes in particular locations.

While these gestures could seem particularly unsafe to be performed while driving in nowadays cars, semiautonomous cars in the future could allow for such “unsafe” but simple gestures when the car is undertaking the driving task from the user.

The think-aloud protocol used during the elicitation study revealed that most gestures have been suggested just because they seemed intuitive and easy to perform, as Morris et al. have found [144]. Nevertheless, four users explicated that they performed gestures similar to those of a smartphone or tablet. Another participant, proposed tap gestures to replicate the same interface of a common MP3 player (iPod). This insight confirms Valdes et al.’s user elicitation results [209] and Jetter et al.’s [100] blended interaction theory, where the knowledge about existing interfaces can have a strong influence on the expected behavior of new interfaces.

7.2.6 WheelSense V4 and Comparison with Other Modalities[25]

7.2.6.1 Explored TGIF Aspects

As for other WheelSense systems, also this system allowed to explore the use of tangible gestures for the *control* application domain, specifically, for controlling the infotainment system of the car. Differently from previous systems, here the interaction was designed with the *user-driven* method, through a gesture elicitation (described in Section 7.2.5) conducted with 20 participants (Group 1), using the Logitech G27 steering wheel as object on which gestures should be performed. In this case, the *hold* component was not a given requirement. Elicited gestures still exploited the *form* affordances provided by this object, but in this case, the steering wheel form was exploited more as a spatial reference rather than for grasping affordances. Taking into account the type of elicited gestures, I have chosen to use an *embedded and embodied* technological approach, using capacitive sensors. Software-side, *heuristics* have been used to recognize gestures. The design space of TGIF explored through this system is represented in Figure A5 in Appendix A.

7.2.6.2 Design Rationale and Methods for WheelSense V4 and the Compared Systems

In section 5.1, I have presented two different methods for designing tangible gestures: expert-driven or user-driven. In Sections 7.2.2, 7.2.3 and 7.2.4, I presented three tangible gesture interactive systems for interacting with the IVIS where tangible gestures were designed by experts. Here, I present a new WheelSense system were tangible gestures were designed with a user-driven approach, i.e., based on the elicitation study presented in Section 7.2.5. More specifically, the six most popular gestures proposed by Group 1 for each of the six commands (Volume Up, Volume Down, Next, Previous, Select

[25] This work has been pushed in 2016 in [6] in collaboration with Stefano Carrino, Francesco Carrino, Maurizio Caon and the department of Psychology of the University of Fribourg.

and Back.) have been adopted as gesture set for this system. In this case, the technology to recognize gestures has been chosen after the definition of the gesture set, according to the peculiarity of the gestures proposed by users. In previous systems, instead, the prior knowledge of the technology used to recognize gestures has influenced the gesture design.

In particular, the main goal of this study was also to compare the use of tangible gestures as interaction modality for IVIS control (WheelSense V4) with two popular interaction modalities already used in modern cars, i.e., touch and speech. The aim was to compare the three interaction modalities in terms of driving performances and user experience. Since tangible gestures on the steering wheel have been recently introduced as interaction means for the IVIS, no previous comparison exist with the other two more common interaction modalities: speech and touch. In particular, the speech interface was designed to comply with the "eyes on the road, hands on the steering wheel" approach, combined with a Head-Up Display (HUD). The tangible gesture interface was designed on purpose for this study, following the aforementioned *user-driven* design, and was also combined with a HUD. The third interface used a popular interaction modality in modern cars, i.e., a touch interface on the central dashboard, which does not comply at all with the "eyes on the road, hands on the steering wheel" approach, since the user has to leave one hand from the steering wheel and to look to the central dashboard in order to interact with the IVIS. Following these considerations, it could be expected a negative impact of the touch interface on driving performance. Besides driving performance, the user experience of the driver was also investigated; therefore, subjective evaluations on the usability, emotional response and perceived workload were conducted. Since the aim of this study is comparing the interaction modalities, not the interfaces, the rationale behind the interface was designed to be exactly the same. A hierarchical menu interface was adopted and the structure and the elements of the menu were the same for the three modalities. In particular, the goal of the interaction scenario was to navigate the menu in order to accomplish some predetermined tasks. Since the three interfaces shared the same menu structure, the number of interactions required to complete the proposed tasks was exactly the same in the three cases.

7.2.6.3 Gesture Syntax and Semantics for WheelSense V4

In the study presented in Section 7.2.5, 20 users were asked to elicit pairs of gestures to operate an IVIS displayed in a HUD with the six aforementioned commands (Volume Up/Down, Next/Previous, Select/Back).

The gestures preferred by the participants (i.e., elicited more often) were adopted for designing the interaction of the WheelSense V4 system. The selected gestures were the followings:

- Volume UP: hand swipe (toward the top) on the external part of the right side of the steering wheel (touch+move components).
- Volume Down: hand swipe (toward the bottom) on the external part of the right side of the steering wheel (touch+move components)
- Previous: hand tap on the external part of the left side of the steering wheel (touch component)
- Next: hand tap on the external part of the right side of the steering wheel (touch component)
- Select: hand tap on the frontal part of the top of the steering wheel (touch component)
- Back: hand tap on the frontal part of the bottom of the steering wheel (touch component)

The 6 gestures are shown in Figure 44. All the six gestures required to remove temporarily one hand from the steering wheel; therefore, no hold component was present in the gestures, differently from gestures designed for the previous three WheelSense systems. Gestures on the left/right side of the steering wheel were performed with the corresponding hand. For the gestures on the top and the bottom, the users could use their preferred hand according to the situation and personal preferences.

Feedback design is discussed in the next section, since it has been designed to be the same also for the other interaction modalities.

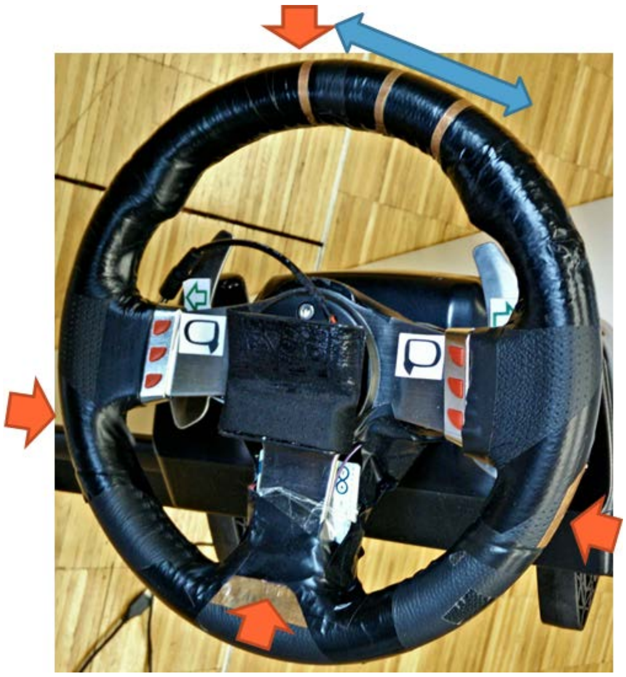

Figure 44. Logitech G27 Steering Wheel equipped with capacitive sensors, also called WheelSense. Red arrows depict hand-taps and the blue arrow depicts the hand-swipe.

7.2.6.4 Design and Implementation of WheelSense V4 and the Compared Systems

The three IVIS interfaces have been designed on purpose for this study, according to the design rationale discussed in the previous sub-section. As anticipated in Section 7.2.6.2 the same menu structure was used for the three interfaces. The first level of the menu contains three items (Music, Contacts and Reminders), which correspond to the three functions implemented in the IVIS (Music Player, Phone Calls and Vocal Assistant). The second level displays the content of each menu. Each content list has a circular structure (the last item is linked with the first one).

In the HUD system, the menu was implemented with a layout similar to the Cover Wheel Layout (but with no 3D effect) [140], as shown in Figure 45. As shown by Mitsopoulos-Rubens et al.[140], this layout structure should not have an impact on performance compared to a list layout, often used in other HUD interfaces. The items of the list are displayed horizontally, with the selected item in the center, highlighted. The HUD was implemented through a software overlay on top of the driving simulator application. The HUD was implemented only for the gesture and speech modalities; for the touch modality, the user relied only on the touchscreen visual feedback, which is the typical configuration available in commercial vehicles. The same audio feedback (music, traffic alerts or simulated phone calls) was provided for the three interfaces through stereo speakers.

Figure 45. Head-Up Display Interface in a music player item

The six aforementioned commands allowed the navigation through the menu items. In particular Volume Up/Down allowed to set the desired volume (in ten steps), the Next and Previous commands allowed going to next and previous elements in the list, while Select and Back were used respectively to select or play an item and go back to the parent item. Each interface implemented the commands exploiting the corresponding interaction modality.

The tangible gesture interface has been designed to comply with the gesture set described in Section 7.2.6.3. In order to recognize the six tangible gestures, a new WheelSense system (WheelSense V4) has been developed. The system adopts an *embedded and embodied* approach and is based on capacitive electrodes integrated in a Logitech G27 steering wheel (depicted in Figure 44). The Freescale MPR121 capacitive sensor is used to measure the capacitive values from the six electrodes, which are collected by an Arduino Uno board. The electrodes were built sticking strips of copper tape on the steering wheel. An important integration work has been done to mask the wires of the electrodes from the user sight as well as to shield the wires from capacitive interferences. Black electrical tape has been used to obtain a soft and smooth surface that ensured a good comfort to the participants while driving and gesturing on the steering wheel. To avoid encumbering cables, the sensor data were sent through a Bluetooth connection to the PC where a C# application recognized gestures through simple *heuristic* algorithms. Therefore, the system did not require to be trained with each user. The system did not require explicit gesture segmentation, which was performed automatically by the system, thanks to the particular locations chosen for gestures. To avoid undesirable false positives, the gesture recognizer was automatically disabled by the system while steering right or left (thanks to an integrated accelerometer) and was manually disabled by the experiment investigator while no task was assigned to the participant.

For the vocal interface, the vocal commands associated to the six functions were simplified as much as possible. The following list reports the six vocal commands:

- Volume Up: “Up”
- Volume Down: “Down”
- Previous: “Left”
- Next: “Right”
- Select: “Select”
- Back: “Back”

During the experiment, the vocal commands were recognized through a Wizard of Oz approach [111]. This method allowed creating the sense of a natural interface where the time between the user’s command and the system execution was similar to the dialog between humans. To keep the interface as natural as possible, the user did not require to pronounce additional vocal keywords (such as “OK Google”) or to press a button to segment the vocal commands.

The touch interface was designed for an Android Tablet with a screen size of 10.1 inches (Samsung Nexus 10). The whole surface of the screen was exploited to make the items of the menu as big as possible, in order to facilitate the touch interaction and the readability of the information. Each button represented a command. Simple icons were used to simplify the interaction with the interface. Figure 47 shows the placement of tablet and provides a hint of the user interface. Figure 46 shows four views of the interface.

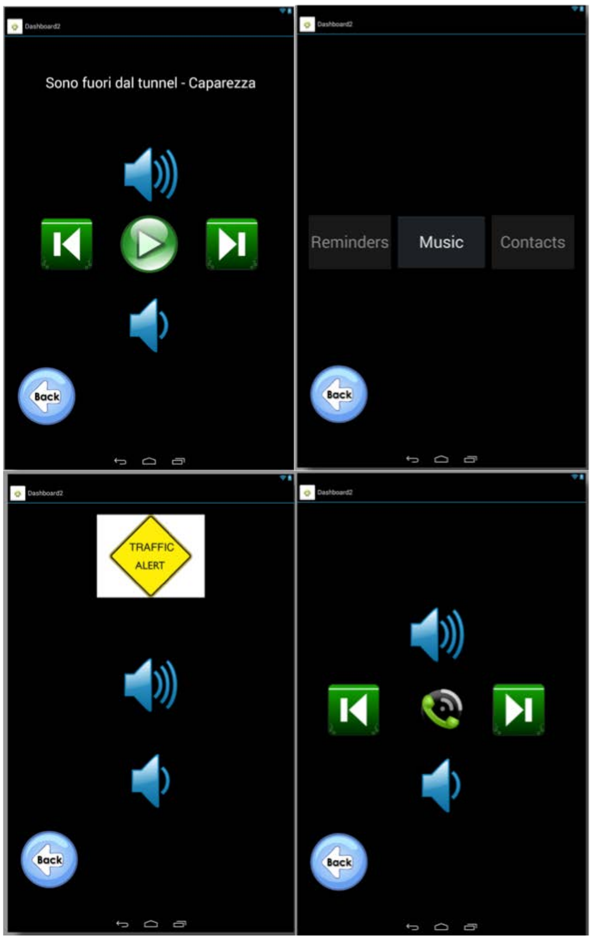

Figure 46. Four views of the touch interface.

A realistic driving simulator was running on an Intel Core i7 Windows PC and displayed on three 19” 1280x1024 LCD screens. Car and IVIS audio feedback was provided through a pair of stereo speakers positioned in front of the user. A Logitech G27 set, including the steering wheel, pedals and the gear stick, was used to control the vehicle in the driving simulator. The tablet was positioned on a reclined portrait stand to simulate a touch interface of the dashboard. The position of the tablet was easy to reach; however, in order to interact with the system, the participants had to glance at the screen, remove one hand from the steering wheel and touch the desired button. The experimental setup is shown in Figure 47. All the interactions with the IVIS were logged by the system.

Figure 47. Experimental setup

7.2.6.5 Evaluation

In order to measure the impact of the interaction modalities on the driving performance as well as on the driver experience, the test scenario should be as close as possible to a real driving scenario. Therefore, a realistic driving simulator has been chosen to ensure obtaining a measurement of perceived cognitive workload and of the emotional response as close as possible to real driving conditions. Although standard measurement tools for assessing driving performances, such as the Lane Change Task [135] or for assessing the cognitive workload, such as the Detection Response Task (ISO/CD 17488), are generally preferable, they alter profoundly the driving experience and therefore they have not been adopted in this study.

In a one-factorial design, IVIS interface was used as a between-subjects variable, being varied on three levels: participants used either a touch interface, or a speech-based interface or a gesture interface. Participants had to complete different tasks in a driving simulation. The dual task paradigm was applied. The primary task was to drive a defined route in the driving simulation. Participants were asked to respect traffic rules and to drive as safe as possible without stopping on the route. To follow the right itinerary participants were guided via verbal instructions by the investigator. The secondary tasks consisted of different interactions with the IVIS that had to be executed while driving. They were inspired by real tasks typical in IVIS usage. On defined positions of the itinerary, pre-recorded verbal instructions were presented to the participants, describing what they had to do. The following tasks had to be completed during the drive:

- Lower the volume of an incoming traffic alert
- Search for a specific person in the contacts list and start a phone call
- Five tasks asked the participants to do some changes in a media player, such as playing a specific song or adjusting the volume

In order to assess the driving performance and the driver's experience with the IVIS, several elements have been measured:

Primary task performance: three measures were recorded: (1) driving time: the time participants needed to drive the route, (2) violations: the number of infractions of the traffic rules, (3) accidents: number of collisions while driving the route. Violations and accidents are automatically recorded by the driving simulator.

Secondary task performance: the number of interactions and the completion time for each task was recorded.

Subjective workload: the Driving Activity Load Index (DALI) was applied [154]. The DALI is a modified version of the National Aeronautics and Space Administration-Task Load Index (NASA-TLX), which has been especially adapted to the driving context [85]. On a seven-point Likert scale (1 = low; 7 = high) participants rated their perceived workload on seven factors (visual demand, auditory demand, tactile demand, temporal demand, interference, attention and situational stress).

Perceived usability and learnability: the System Usability Scale (SUS)[31] was applied after system usage. The SUS consists of 10 items and is well suited for a quick comparison between systems in regard of usability. Items were rated on a seven-point Likert scale (1 = strongly agree; 7 = strongly disagree), which results in a score between 0 and 100. The overall internal consistency of the questionnaire (Cronbach's $\alpha > .91$) is high. The learnability and usability factors have been measured using two and eight items of the SUS questionnaire as explained by Lewis and Sauro in [127].

Emotions: the PANAVA-KS (short scale for the assessment of positive affect, negative affect and valence) [181] was applied. It consists of 10 bipolar adjectives that describe different affective states. Items were rated on a seven-point Likert scale with the two extreme points (e.g. happy: very happy, very unhappy). The psychometric properties of the questionnaire are sufficient (Cronbach's α =.83 for positive affect, .76 for negative affect and .74 for Valence).

For measures of user performance and subjective user ratings, a one-factorial analysis of variance (ANOVA) was carried out, followed by explorative post-hoc comparisons, for which a Bonferroni correction was applied. For the analysis of the data on participants' affective states, a one-factorial analysis of covariance (ANCOVA) was carried out, with the initial baseline measure (taken prior to task completion) used as covariate.

The sample of this study consists of 60 participants (83.33% female). All of them were members of the University of Fribourg (Switzerland), in possession of a valid driving license and aged between 18 and 40 years ($M = 23.22$). Self-reports collected with a questionnaire before the experiment showed that the participants rated themselves as average drivers ($M = 4.37$ on a seven-point Likert scale from "beginner" to "expert"). With regard to simulator experience, they rather rated themselves as beginners ($M = 2.30$). When asked about the frequency of usage of different interaction modalities in the car in a 7 point Likert scale, buttons were by far the most used interface (M=4.6 SD=1.9) followed by touchscreen (M=2.0 SD =1.6) and speech interaction (M=1.2 SD=0.7). Concerning the interaction modality that users considered as easier to use to interact with an IVIS, 5 participants answered the speech interaction, 1 the touch interaction and 54 an interface with buttons. Since there are no commercial cars equipped with a gesture interface, it was not included as an option in the questionnaire.

The data of the driving performance are presented in Table 4. Data analysis indicated no significant difference between the different interaction modalities for driving time, driving errors and accidents (all $Fs < 1$).

Table 4. Performance values (lower is better) on the driving task as a function of the interaction modality

	Touch control	**Voice control**	**WheelSense**
	M (SD)	M (SD)	M (SD)
Time (s)	917 (124)	935 (154)	902 (101)
Driving errors (#)	34 (10)	34 (13)	36 (11)
Accidents(#)	2.8 (2.3)	3.6 (3.1)	3.0 (2.7)

Task completion time. The analysis of the data of task completion time (see Table 5) revealed a significant effect between the different interaction modalities (F = 5.3; df = 2, 57; p < .01), with post-hoc analysis indicating that task completion time in the touch modality was lower (i.e., higher performance) compared to the WheelSense modality. The other post-hoc comparisons did not reach significance level.

Interaction efficiency. The data about the number of user interactions are also presented in Table 5. Statistical analysis (F = 16.8; df = 2, 57; p < .001) revealed a lower number of user interactions (i.e., higher performance) for the speech control condition compared to the other two conditions (large effects with r between .63 and .68). Since the menu structure for the three IVIS was the same, those findings indicate that participants committed fewer errors in the speech condition.

Table 5. Performance values on the secondary tasks (lower is better) and results of post-hoc tests (Bonferroni corrected) as a function of interaction modality.

	Touch control	**Voice control**	**Wheel Sense**		
	M (SD)	M (SD)	M (SD)	*p*	*r*
Task completion time (s)	34.8 (8.3)	39.5 (9.6)	43.5 (7.3)		
	↑		↑	< .01	.49
Interaction efficiency (#)	15.4 (2.0)	12.9 (0.9)	15.1 (1.4)		
	↑	↑		< .001	.63
		↑	↑	< .001	.68

The data of mental workload, perceived usability and learnability are presented in Table 6. Mental workload is reported as the average of the seven items of the Dali questionnaire. As expected, the perceived workload is above average for all the three interfaces. The SUS scores for the touch and voice modality are slightly above average (68 is the average for a SUS score) while the Wheel Sense interface scored slightly below the average. However, the calculated ANOVAs showed no significant effect of interaction modality on the different subjective evaluations of the IVIS systems (all three Fs < 1).

Table 6. Subjective evaluation of the IVIS systems (higher is better) as a function of the interaction modality.

	Touch control	**Voice control**	**Wheel Sense**
	M (SD)	M (SD)	M (SD)
Mental workload (1-7)	5.3 (.56)	5.0 (.71)	5.1 (.98)
Perceived Usability (SUS) (1-100)	72.7 (16.0)	72.1 (16.5)	66.2 (15.4)
Usability (1-100)	69.0 (18.6)	69.8 (16.6)	64.6 (14.7)
Learnability (1-100)	86.6 (15.6)	80.0 (23.3)	73.7 (28.9)

Table 7 presents the summative results of the questions relative to valence, positive affect (PA) and negative affect (NA) in the PANAVA-KS questionnaire (two, four and four questions, respectively). For the three interfaces, both valence and positive affect are slightly above average, while negative affect is around average. Statistical analysis revealed no significant influence of interaction modality on any measure of user affect (FValence = 2.3; df = 2, 54; p > .05; FPA < 1; FNA = 2.8; df = 2, 54; p = .07). The covariates initial PA showed a significant effect on users' PA ratings after task completion (F = 8.0; df = 1, 54; p < .01) as well as the covariate initial NA was linked with participants' NA ratings (F = 15.3; df = 1, 54; p < .001). No other effect of a covariate reached significance level.

Table 7. Ratings of users' affective states after task completion as a function of the interaction modality.

	Touch control	**Voice control**	**Wheel Sense**
	M (SD)	M (SD)	M (SD)
Valence (2-14)	10.4 (2.2)	9.1 (1.8)	10.1 (2.0)
Positive Affect (4-28)	19.2 (3.0)	18.8 (2.5)	18.7 (3.6)
Negative Affect (4-28)	17.0 (3.4)	17.2 (2.8)	15.3 (3.4)

The main goal of this study was to evaluate three different interaction modalities of in-vehicle information systems in the context of a driving simulation study. Measures of driving performance, interaction performance, workload, perceived usability and user affect were recorded. The main results showed some advantages of the speech control (with regard to interaction efficiency) and touch control (regarding task completion time) for performance on the secondary task. This could be explained because the users could rapidly execute the same command on a touch interface, especially if they have to press the same button. Since the interaction with the touchscreen is visually demanding, users were sometimes imprecise, making mistakes, for example going too far in the item list in respect to the desired item. The same considerations could not be done for the gesture interface on the steering wheel, probably because the users were not acquainted with this interaction modality.

The interaction modality of the IVIS had no effect on all the other measures assessed in this study (e.g. driving performance, subjective evaluation of the IVIS system as well as on the users' affective state). This result is rather surprising since it was expected that the use of interaction modalities allowing the drivers to keep their eyes on the road (i.e., speech and gestures) would result in increased driving performance (e.g., decreasing driving errors and accidents) compared to the use of the touch modality, for which the driver needs to shift her attention from the road towards the central dashboard.

Gestures and speech are very innovative and uncommon systems. Users did not have previous experience with similar systems; therefore, it is surprising and satisfactory that ratings of their usability did not differ from the ratings of the better-known touch system. Indeed, there was a remarkable difference between the learnability score of the touch and gesture systems, although not statistically significant ($p=0.084$).

Concerning the absence of difference in driving performance and subjective workload, it can be argued that the novelty of speech control and gesture control might have had some negative influence on performance and perceived workload. Since the interaction with these new systems was uncommon and novel, this might have been the reason for a decrease in driving performance and an increase in subjective workload. The advantages of those systems in comparison with the touch system due to the eyes on the road-principle might hence have been compensated by their disadvantages due to their novelty.

The study was intended to measure several aspects of different interaction modalities, resulting in a complex experiment that has unavoidably some limitations. In this study, the three interfaces were designed with the same rationale. However, each interaction modality could offer several advantages that were not exploited in the designed interface for the sake of a “fair” comparison. For example, considering a speech interface, it could be possible to issue directly an order (e.g., “System, call Home” or “System, play The Man Who Sold The World”). This will reduce the time in which the driver’s focus switches from the primary to the secondary task, and will reduce the overall time spent interacting with the IVIS. Nevertheless, a segmentation approach is generally required for most speech recognizers: the current commercial systems require the user to press a button or to use a keyword to activate the vocal interface (such as the word “System” in our example or “Ok, Google” and “Hey Siri” in other systems based on speech recognition). In this study, this aspect was overlooked, using a “Wizard of Oz” approach and considering the system as able of automatically segment the vocal commands. Finally, the design of interaction-specific solutions can be extremely varied. Concerning the gesture interaction, the chosen gesture vocabulary could have an impact on the perceived usability and emotional response of the users, although the gesture elicitation approach should have optimized this aspect. The number of interaction errors made by the users had not been counted, because it was difficult to discriminate user errors from errors of the recognition system (especially in the Wizard of Oz scenario). Finally, concerning the driving simulator, the real case scenario should have allowed participants to experience a workload typical of a city center driving scenario. Nevertheless, simulated driving cannot be compared with a real driving experiment, also because participants did not have time to get completely acquainted with the driving simulator (participants rated themselves as beginners and have only ten minutes to learn the driving simulator and the IVIS interface). For this reason, the number of driving accidents and violations were sensibly higher than what could be expected in a naturalistic driving scenario.

7.3 ADA Lamp[26]

7.3.1 Introduction and Field Requirements

In this Section, I describe the work done in the domain of interpersonal communication at distance and emotional interfaces, and in particular the design and development of the ADA lamp system. The fields requirements have been studied through a series of preliminary works in which I investigated the possibility to produce and deliver emotionally-rich messages in a more natural way [37]. A preliminary study for sharing emotions in social awareness streams through natural facial expressions has been

[26] This work has been published in 2014 in [4] and in 2015 in [8]. The paper presented in [4] won the best work-in-progress award during the IHM’14 conference.

conducted in order to understand if users prefer natural means for sharing emotions instead of emoticons [38]. In the proposed system, messages and emotions are shared through the Twitter social awareness stream. As input, facial expressions are recognized either by a mobile device, thanks to the frontal camera of the smartphone, or by the environment, i.e., with a depth camera associated to a smart TV. In order to share emotions, the user has to mimic the corresponding facial expression. As output, a robotic painting mimics the emotions shared by the people that the user is following and a RGB lamp matches the color of the ambient light with the valence of the shared emotions. The system was demoed at the ACII 2013 conference, where 85 attendant passed by the stand and 29 stopped interacting with the system. People that tested this preliminary system appreciated the proposed forms of interaction for sharing emotions, although some of the attendants felt that acting facial expression in front of a camera in order to send an emoticon was unnatural and not appropriate for a public space [36]. Concerning the system output, they particular appreciated the display of facial expressions in the robotic painting, thanks to the life-like behavior of the object, while the ambient colored light was barely noticed by the attendants.

These preliminary observations led to the idea of integrating the facial expression representations inside the RGB lamp, animating the lamp with a life-like behavior, allowing users to interact through gestures performed on the lamp instead that using acted facial expressions. Therefore, the purpose of this study was to create a device able to support meaningful emotionally rich interactions and to support these interactions on a long-term basis.

7.3.2 Explored TGIF Aspects

This system allowed exploring the *affective communication* application domain. The interaction was designed using the *expert-driven* method, with a *reification* of new gestures performed on a given object, the lamp. Following field requirements, all gestures were designed including the *touch* component, to support the emotionality of the interface. Gestures were designed exploiting the *form* affordances provided by this object and the *anthropomorphic* affordances supported by the internal facial expression display. The interaction design explored also the possibility to animate the object with a *life-like* behavior to support long-term interactions. The resulting semantic construct used for associate meaning to gesture was the *noun+verb* metaphor. The technological approach chosen for the implementation of the system was *embodied and embedded*, with transparent capacitive electrodes used to detect the tangible gestures. Software-side, *heuristics* have been used to recognize gestures. The design space of TGIF explored through this system is represented in Figure A6 in Appendix A.

7.3.3 Design rationale and methods

The system presented in this Section has been designed following the insights collected during the preliminary studies presented in Section 7.3.1, the insights collected in the literature analysis presented in Section 3.4 and the guidelines and common practices presented in Section 5.4 and 5.5. In particular, from the literature analysis, I took inspiration from the concepts for life-like interactive objects proposed by Schmitz [184]. Indeed, the Anthropomorphic Display of Affection (ADA) [8] is a system designed to investigate tangible gestures as an emotional communication means. The aim is to provide the user with tangible affordances that can facilitate the interaction with an artifact (a spherical lamp), in particular for performing gestures that carry strong emotional contents. To this purpose, the lamp should be able to represent emotional states (through facial expressions) and to interpret emotional gestures.

The ADA lamp can be used in two different scenarios: as a smart companion and as interface for affective interaction in long distance relationships. In the first scenario, the anthropomorphic lamp presents personal intelligence and behavior; it reacts to the users' gestures and tries to attract their attention showing unpredictable behaviors. These latter can create a sense of surprise that can sustain the user engagement on the long term [184]. An advanced version of the lamp could adapt its behavior to the current mood of the user. The companion lamp could be used just for entertainment or to share emotional states on the social networks, of course, in addition to illuminating the surrounding environment.

In the second scenario, the lamp can function as an interface for emotional telepresence and affective computer-mediated communication. In fact, ADA will represent the emotional states of a distant loved person. The emotional state can be shared via private messages on social networks or interacting directly with the lamp. The emotional state of the lamp can also change according to the gesture performed from the distant person on the other lamp: a gesture with positive valence will make the lamp to go to a positive emotional state and vice-versa. In this way, the separated people can naturally interact directly through a couple of ADA lamps limiting the usage of other less natural digital communication means.

7.3.4 Interaction Design: Tangible Gesture Syntax and Semantics, Feedback

The interaction has been designed to exploit the anthropomorphic affordances provided by the spherical *form* of the lamp and the facial expressions that the lamp can mimic (a mouth that smiles or is frown and eyes that can blink). These affordances help the user recognizing the metaphor of a human head. These life-like affordances aim at facilitating the discovery and learning of the gestures that the user can perform on the lamp and that the system is able to recognize. Indeed, the tangible gestures recognized by the lamp are those typically used in human-to-human communication, such as caressing, kissing or slapping.

In particular, five gestures have been designed for the ADA lamp: the users can caress (touch+move), hug (hold+touch) or kiss (touch) the head to convey positive emotions, but they can also slap (touch(+move)[27]) and shake (Hold +move) the head to convey negative emotions (Figure 48).

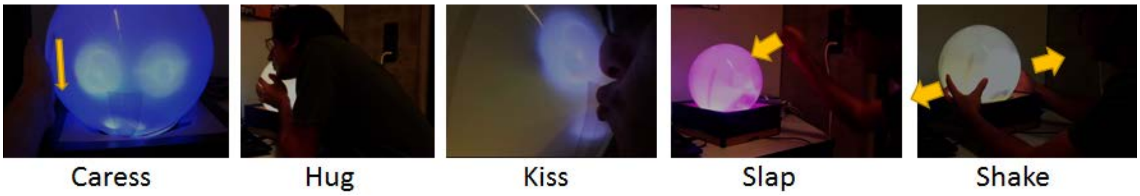

Figure 48. The five gestures designed for the ADA lamp.

All gestures that can be performed on the ADA lamp include the touch or hold component. The main reason of this choice is the importance of touch in social relationships, as anticipated in Section 4.3 and discusses in Section 5.4. The semantics of tangible gestures is already known by the user, since gestures have been designed to be the same than those used in interpersonal face-to-face communication. The affordances provided by the object, i.e., the rounded shape with projected eyes and mouth, helps the user guessing the types of tangible gestures that can be used to interact with the system. Since the object has a clear reference to a human head and the gestures are typically used in human-human interaction, the semantic classification of these tangible gestures falls in the category of the noun+verb metaphor.

As feedback, the lamp should be able to represent different emotions (happy, sad, trusty, angry and ecstasy) through facial expression (frown or smile, eye blink). To increase the effectiveness of emotional communication, the facial expressions of the lamp have been also mapped to a RGB color scheme based on the Plutchick’s wheel of emotions [161] (cf. Figure 49 and Figure 51).

7.3.5 Implementation

The system has been developed according to the embedded and embodied approach, modifying a real-word object that is commonly used during our daily routine: a table lamp. The first phase of the

[27] The slap adds an additional move component to the touch gesture since the effect (which is measured also from the system) is to suddenly make move the lamp. This particular interpretation of the touch+move gesture should be interpreted under an object-centered point of view of the TGIF taxonomy, as discussed in Section 8.2.2 and 9.2.

development was focused on giving to the lamp an anthropomorphic form. A commercial spherical table lamp fitted perfectly the purpose of obtaining a form similar to a human head. The light bulb has been replaced with a strip with 60 RGB addressable LEDs disposed on a cylindrical structure. An internal separation with paper and plastic structures conveys the light of specific LEDs for the left and right eyes and for a frown and a smiling mouth. LED colors and intensities are individually controlled through an Arduino Uno board (Figure 50). By selectively lighting up the different zones with different colors, it is possible to choose an arbitrary color for each face region, to hide one of the two mouths or simulate an eye blink. The five different facial expressions are shown in Figure 49.

Figure 49. The five emotional states of the ADA lamp: happy, sad, trusty (eye blink) (2x), ecstasy (2x) and angry.

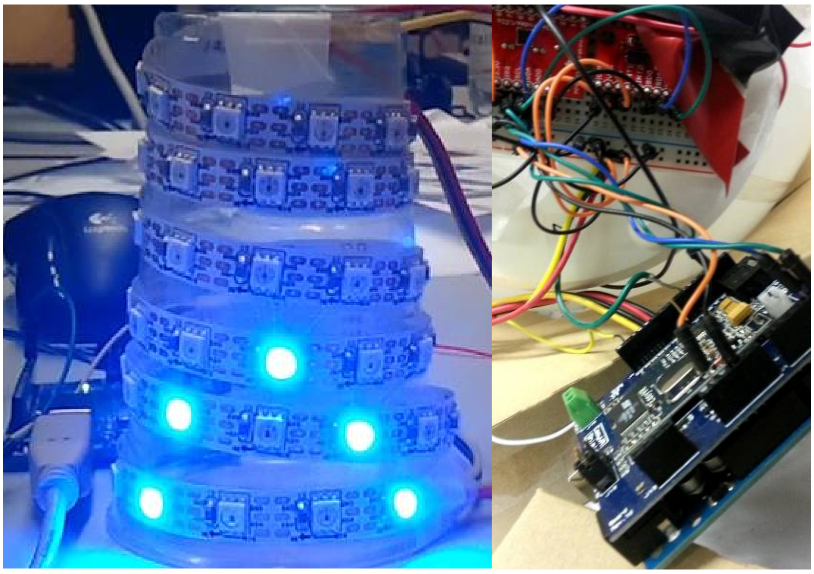

Figure 50. On the left, the cylindrical structure of the LED strip, displaying a frown mouth; on the right, the Arduino Uno with the EasyVR shield and the MPR121 and ADXL345 sensors.

As a second step, I implemented the recognition of gestures performed on the surface of the ADA lamp. A MPR121 capacitive sensor has been connected to the Arduino Uno board and six transparent Indium Tin Oxide (ITO) electrodes have been attached all around the lamp surface and connected to the MPR121. Caressing, hugging and kissing, which rely mostly on the touch component, are recognized through a simple threshold-based algorithm that check the contact on the different zones covered by the electrodes. Slapping and shaking, which imply a movement of the object, are recognized through an ADXL345 3-axis accelerometer, also connected to the Arduino.

Finally, the life-like behavior of the companion lamp has been implemented with a state diagram. The states in the diagram in Figure 51 represents the 5 different emotional states of the lamp, with their corresponding colors and facial expressions. Transitions to another state can be obtained by performing on the lamp a gesture, a sequence of gestures, or by a period of inactivity, i.e., absence of any gesture. In Figure 51, I show an implementation of the state diagram for the companion scenario of the lamp. In order to improve clarity, I illustrated separately the transition diagram for negative valence gestures and inactivity (on the left) and the transition diagram for positive valence gestures (on the right). Multiple arrows of the same color depict gestures that need to be performed multiple times in order to activate the transition. The transition diagram and the gesture recognition algorithms are executed in a C# application running in a PC. The ADA lamp is connected to the PC through a USB serial connection. The state diagram can be easily customized to obtain different behaviors of the lamp. For example, following the examples provided by the provoking article of Buttrick et al.[34], one could imagine that the lamp reacts in a positive manner to gesture with a negative valence and, vice versa, gets angry when the user performs gestures with positive valence. The state diagram can have input also from other stimuli, for example inferring the mood of the user from Twitter or from life-logging devices.

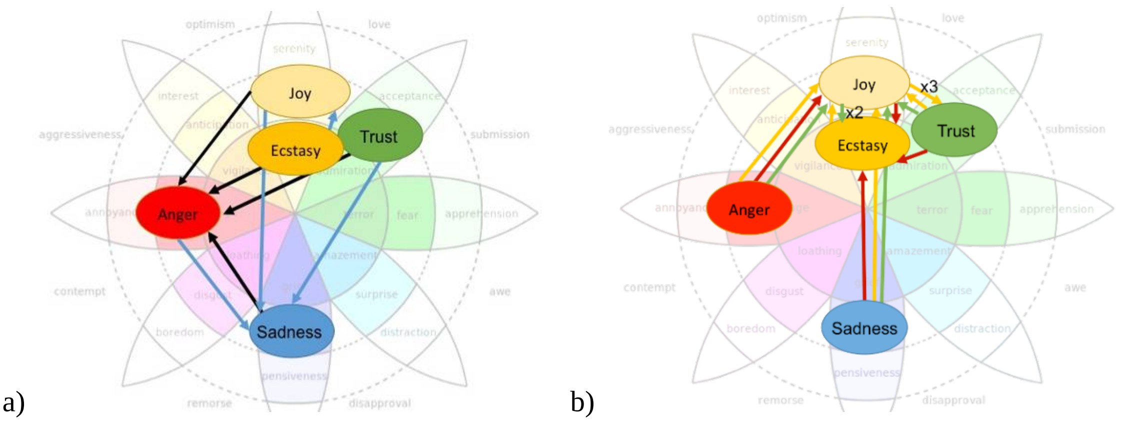

Figure 51. On the left, the diagram for the negative valence gestures; b) on the right, the diagram for positive valence gestures.

The system is still in a prototype state, since several elements should be improved before conducting a long-term evaluation of the system. Although the ITO transparent electrodes do not hinder the sight of the underlying facial expressions of the lamp, they still limit the smoothness of the surface while performing touch+move gestures and limit the ability to recognize gestures to the small zones of the lamp where they are attached. Ideally, a full spherical touch grid should be implemented in order to improve the ability to recognize gestures. This is possible only through expensive ad-hoc capacitive touch solutions or through Frustrated Total Internal Reflection (FTIR) of infrared light, a solution typically used on tabletops. Moreover, the solution for shaping the projections of the RGB LED light based on custom-built shapes, although able to produce vivid colors, is not durable neither easy to reproduce in large scale. Therefore, it should be improved or substituted with a spherical projection. Both FTIR and spherical projection solutions have been shown in [26].

Finally, in order to enrich the interaction modalities of the lamp, it could be considered the implementation of haptic feedback through a vibration motor and audio feedback through an Arduino EasyVR shield. Using this module, it is also possible to implement vocal interactions, such as short phrases or the names of people to whom the emotional gestures are devoted.

7.3.6 Evaluation

During two interactive demonstrations of the ADA lamp at the IHM'14 and TEI'15 conferences, I observed several participants interacting with the lamp, which was programmed to react as a smart companion. The anthropomorphic affordances stimulated conference attendees to interact with the lamp, although I noticed that most caresses were performed on the top of the lamp while the system was designed to recognize only caresses performed on the side, as one would perform on an adult. I suppose that this behavior is due to the position of the lamp (on a table), which is lower than an adult human head, thus ergonomically facilitating gestures on the top of the lamp (as one would do on a child or on a pet head). Moreover, people were generally reluctant to slap or shake the lamp unless explicitly encouraged.

An additional experience should be conducted with a more finished prototype in order to investigate the user interactions with the lamp in a longer-term perspective. The second interaction scenario should be also studied more in depth: Hemmert et al. [87] demonstrated that it is important to carefully design devices for emotional telepresence because they can result awkward or disturbing. For this reason, it is crucial to conduct further research on user acceptance and also studying potential cultural differences in the degrees of acceptance.

7.4 Hugginess[28]

7.4.1 Introduction and Field Requirements

The purpose of the Hugginess system was different from all the other systems developed in this thesis. Indeed, while in the other projects the users communicated through tangible gestures merely with the machine, or eventually with another user through a machine, here the purpose is encouraging interpersonal face-to-face communication, in particular through the hugging tangible gesture.

Since hugging is able to elicit positive emotions, it could be used as a therapy [110, 221]. Using hugging as a means to give rise to such positive emotions can be seen in the frame of positive psychology. In a society where technology, and in particular mobile devices are often seen as a barrier for interpersonal face-to-face communication, the purpose of the Hugginess is demonstrating that such technology can also be able to encourage people to have physical contact, and in particular to hug and thus to improve their general well-being. In fact, the effects of hugging can be beneficial to both individuals and society; therefore, as stated by Forsell and Astrom, "encouragement might become a vital part of public wellness programs" [70]. Compared to handshaking, which is another touch gesture commonly used for greeting, hugging is a closer and more affectionate form of hailing. Other studies suggest that hugging behavior seems to create a more emotional quality in the initial phase of a conversation, as compared to a verbal greeting or a handshaking, which leads to a more formal initial phase and conversation [70]. The change from handshaking to hugging during an encounter is associated with a greater emotional involvement [70]. Some people, who have extraverted personality, are naturally facilitated in taking the initiative in hugging, as extraversion is associated with spontaneity and sociability[109]. On the other hand, anxious people, who lack self-esteem and self-confidence, may present a decreased likelihood of taking the initiative to hug [109]; in this case, an external factor that encourages this behavior can be decisive. Moreover, the current trend of the technology-driven evolution of the human communication is separating people one from each other. In fact, Turkle stated that nowadays people prefer to use computer mediated-communication in order to hide from each other but remaining constantly connected [204]. This change deprives the human communication of the interpersonal touch and many people in society today may already be suffering from a shortage of tactile stimulation, a phenomenon that Field calls "touch hunger" [65].

[28] This work has been published in 2014 in [11]. The development of the work has been supported by Lucy Linder.

Within this context, the aim of this system was to recognize and reward the hugging gesture between to people, in order to encourage people to hug more. The reward was conceived as an exchange of information between huggers and, in this case, the interaction should support the intimacy between the huggers and the same time the system should support the security and privacy of the exchanged information.

7.4.2 Explored TGIF Aspects

This system, as the previous system presented in Section 7.3, allowed exploring the *emotional communication* application domain, this time focusing on interpersonal interactions in physical encounters. The interaction was designed using the *expert-driven* method, but modeling and augmenting an existing gesture commonly used by humans for social interactions: hugging. This gesture includes the *hold* and *touch* component to increase the emotionality of the experience. The gesture naturally exploits the *form* and *deformability* affordances provided by the human body, and further augmented the deformability affordance with a soft patch on the t-shirt for detecting the hugging pressure. Since the system for recognizing gesture is a smart t-shirt, the technological approach chosen for the implementation of the system is at the same time *wearable* (since it is worn by the user) and *embodied and embedded*, since it is an augmentation of the object of the tangible gesture, i.e., the human body. Software-side, *heuristics* have been used to recognize gestures. The design space of TGIF explored through this system is represented in Figure A7 in Appendix A.

7.4.3 Design rationale and methods

This project strives to encourage hugging through an innovative system that augments the physical gesture with digital information exchange. Reintroducing the hugging behavior in the interpersonal communication can provide many positive effects on human well-being such as reducing stress and strengthening human relationships. These good consequences could help hopefully the human kind in the difficult path towards happiness, as the positive technology discipline teaches us. Besides the practical benefit of the proposed system, the purpose of this concept is stimulating reflections and help people understand the importance of promoting interpersonal touch in the current society. This is clearly an expert-driven design method, which tries to cope with a problem that often escape from users' daily attention. Therefore, the design of the interaction has been focused on enhancing and exaggerating the hugging gesture to maximize the positive benefits and make people reflect about them.

In order to augment the hugging gesture, first the system should be able to recognize it. A wearable approach was the natural solution for implementing the recognition of the tangible gesture. From one side, the technology should disappear behind the fabrics of the everyday-life (and of a t-shirt), putting the hugging gesture at the core of the interaction. At the same time, the system should be clearly visible to all potential users and to spectators in order to encourage hugging and stimulate reflection about its importance. To this purpose, two different versions of the system have been conceived; the first version, a demonstrator of the concept for everyday-life utilization, is constituted by two smart t-shirts connected to a smartphone app to track hugging gestures and collect the augmented digital information; the second version, instead, is a smart hoodie that works as a hug-counter, and has been designed for sensitizing people to the importance of interpersonal touch during free-hug events.

7.4.4 Interaction Design: Gesture Syntax and Semantics, Feedback

As for the ADA lamp, one of the aims of the Hugginess system is to recognize gestures that people usually perform to communicate with other people, especially those with a strong emotional component. In this case, the project focuses only on the hug gesture, but instead of gesturing on an anthropomorphic object, this system is conceived to recognize and enhance tangible gestures performed on other people. The syntax and semantics of the hug gesture are particularly rich: the hold component of two people hugging has an implicit association to reciprocal belonging, while the touch component is important from a psychological and also physiological point of view, stimulating the release oxytocin in the huggers' bodies and therefore increasing their well-being. Touch is also localized in both the chests of the huggers and in the backs, where the hands should apply pressure in order to generate a stronger emotional arousal. These three elements (the hold and touch components and the pressure attribute) have been taken into account for the design of the system. It is worth noting that the tangible gesture performed in the real life, i.e., hugging another person, is the core of the interaction and has the same semantic reference in the digital world. However, it has an additional effect in the digital system: it allows sharing contact information (such as the telephone number) through the fabrics of the t-shirts while hugging, which is a strong metaphor for the importance of the physical contact between people. The hugging gesture between two people wearing a preliminary version of the Hugginess t-shirt is shown in Figure 52. Feedback should be unobtrusive, since huggers should be focused on their gesture instead that on system state. A simple RGB LED should be sufficient to display the system state and to confirm that contact information has been shared.

Figure 52. The hug gesture for the Hugginess system.

7.4.5 Implementation

As anticipated in Section 7.4.1, one of the purposes of the Hugginess system is promoting physical interpersonal touch, augmenting the hugging gesture with an exchange of digital information, such as their phone number or a private message that they would like to share with the people they hug. As a concept, the Hugginess t-shirts could modulate the amount of information exchanged according to the length of the hug, but also to the social closeness of the hugged person. For example, a short hug between two unfamiliar people would exchange only the contact name and email of the two people, while a longer hug would exchange also the telephone number and social network contacts. A hug between two people that already hugged in the past, for example two close friends, would exchange more intimate information about their current mood and recent important events of his or her life, rather than the contact information that they already have. For the first proof of concept system, only the phone number and a private message that is recorded in the t-shirt is shared when hugging.

One of the major objectives of the system design is reflecting the user's need of physical contact in the working principle of the system. While most existing projects that exchange information through contact gestures (e.g., handshaking) rely on a wireless communication (cf. Section 3.4 for a review), the t-shirt designed for the Hugginess project needs physical contact with another smart t-shirt in order to exchange digital information. Indeed, the information is exchanged through conductive patches, which act as wires for a point-to-point communication. Such unique feature of the system has an additional advantage: thanks to the intimacy of the Hugginess interface and to the physical communication channel between the two huggers, the user should be reassured about the privacy of his or her data. The risk of someone stealing his or her data is very low, since no man-in-the-middle attack is possible without a physical interposition between the two huggers. Conversely, the man-in-the-middle attack is a typical limitation of wireless transmissions, which can be easily sniffed if not adequately encrypted.

Following the analysis of the hug gesture presented in Section 7.4.4 and in order to ensure the intimacy of the information exchange, the following conditions should be verified in a hug gesture:

- There is physical contact between the two people's chest
- The two heads are at same height (even if the two people have different heights)
- The users' arms are pressing reciprocal backs

The two first conditions are important to locate a contact surface for exchanging data between t-shirts. Since there is a bidirectional flow of information between the two users, two communication channels composed by a transmitter and a receiver are needed. I chose to attach two conductive patches symmetrically, at a fixed distance from the vertical axis and from the shoulders. I made this assumption in order to ensure proper patch contact between people with different body dimensions (males, females, children, adults). Moreover, I chose to position the transmitter on the left side of the t-shirt, just above the heart. This design aims at inspiring people by depicting metaphorically a flow of information from the heart when people are hugging. The third condition ensures high arousal during the hug and unintended information exchanges during occasional contacts in crowded environments: in order to activate the exchange of information, both users need to press the other user's back. As stressed before, the interaction modality of Hugginess ensures data privacy without the risks that unintended people steal users' data. The design of the t-shirt is shown in Fig. 53.

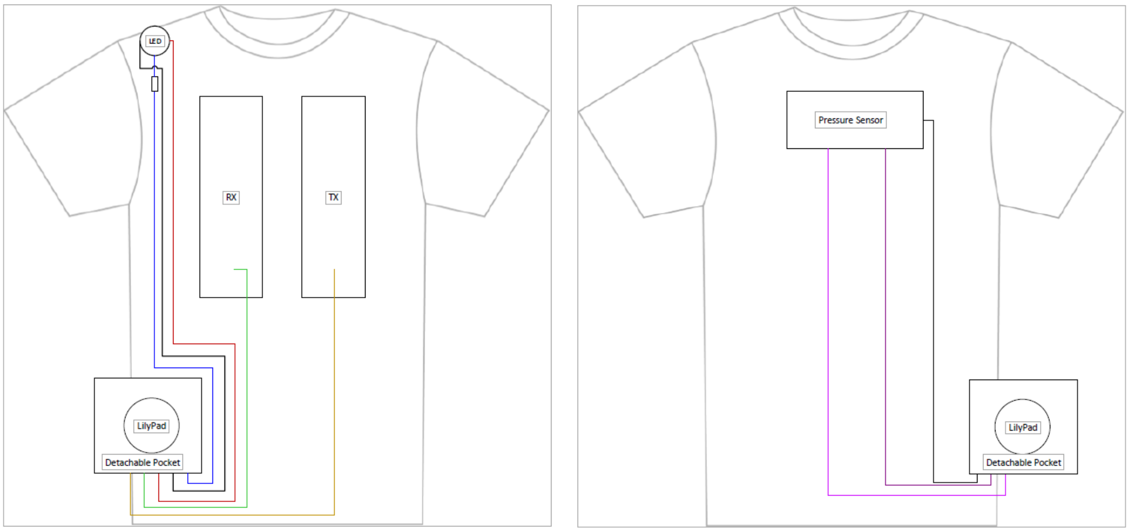

Figure 53. Hugginess t-shirt schematics

The system (Figure 54) has been realized with an Arduino Lilypad, which is sewed to a detachable pocket in order to easily detach the hardware components from the fabrics and wash the remaining of the t-shirt. The Lilypad board is connected to the two conductive patches on the front, one for the transmitter and the other for the receiver, to a pressure sensor realized with a sheet of Velostat piezoelectric material on the back, and to an RGB LED on the shoulder, which shows the status of the information exchange while hugging. Conductive threads have been sewed to the t-shirt to connect the different embedded components. The Hugginess t-shirt is powered by a Lithium battery and can be paired to a smartphone through Bluetooth (a BlueSMiRF module has been added to the LilyPad Arduino).

Figure 54. Hugginess t-shirt prototype

The communication between the two t-shirts has been implemented through the AltSoftSerial library, which emulate a serial port over the two pins connected to the respective TX and RX fabrics. When the two huggers are face-to-face and start hugging, the RX of one hugger get in contact with the TX of the other hugger, and vice versa. As soon as a pressure is detected on the back, data is sent in loop by both t-shirts through the fabrics and as soon as a message is correctly received by the t-shirt (the message integrity is controlled through a checksum), the hug is recognized and the received data are stored in a buffer of the t-shirt.

The t-shirt could work standalone for basic information exchange, without the need of permanent wireless connectivity to the smartphone. In fact, the users can store their static contact information in the t-shirt by connecting it the first time to the dedicated smartphone app. At the same time, the t-shirt can store internally all the information collected during hugs, i.e., contact information of other people and their private message. Nevertheless, the Hugginess app can provide additional ways to encourage people to hug. Indeed, it logs all the user's physical interactions detected by the t-shirt and thus it can infer a lack of physical contact over a long period. Statistics are shown in the application as well as a log of the last hug exchanges. Here the user can also read the private messages shared by the other hugged users through their t-shirt. As an additional encouragement to hug, the smartphone app could be location-aware and could support a proactive behavior to avoid loneliness and lack of physical contact by automatically sending a message to nearby friends, inviting them to go hug the person that is lacking physical contact. Two screenshots of the app are shown in Figure 55.

Figure 55. The Hugginess app

While the Hugginess t-shirt has been developed as a proof of concept for encouraging interpersonal touch during everyday life, a variant of the system has been developed adapting the existing system in order to use it in a Free-Hug event. Therefore, the purpose here was to make the system more attractive and to stimulate people that do not a previous knowledge of the system to hug those that are wearing the system. Since the new system cannot reward the huggers by exchanging contact information (because people that are just passing-by are not wearing a Hugginess t-shirt paired to their smartphone), a new well-visible reward has been designed in the form of a two-digit hug counter. This system has been integrated in a hoodie and the counter has been implemented through two red/green 8x8 LED matrix display. While the conductive patches for digital information exchange have been removed, the smart-hoodie is still able to recognize hug gestures through the pressure sensor on the back and to send the information about the hugs to the paired smartphone. In particular, the app allows counting the number of hugs as well as the time spent hugging in seconds. These two values can also be shown in the two-digit display, switching between them through a parameter of the smartphone app. The Hugginess Hoodies and the relative app are shown in Figure 56.

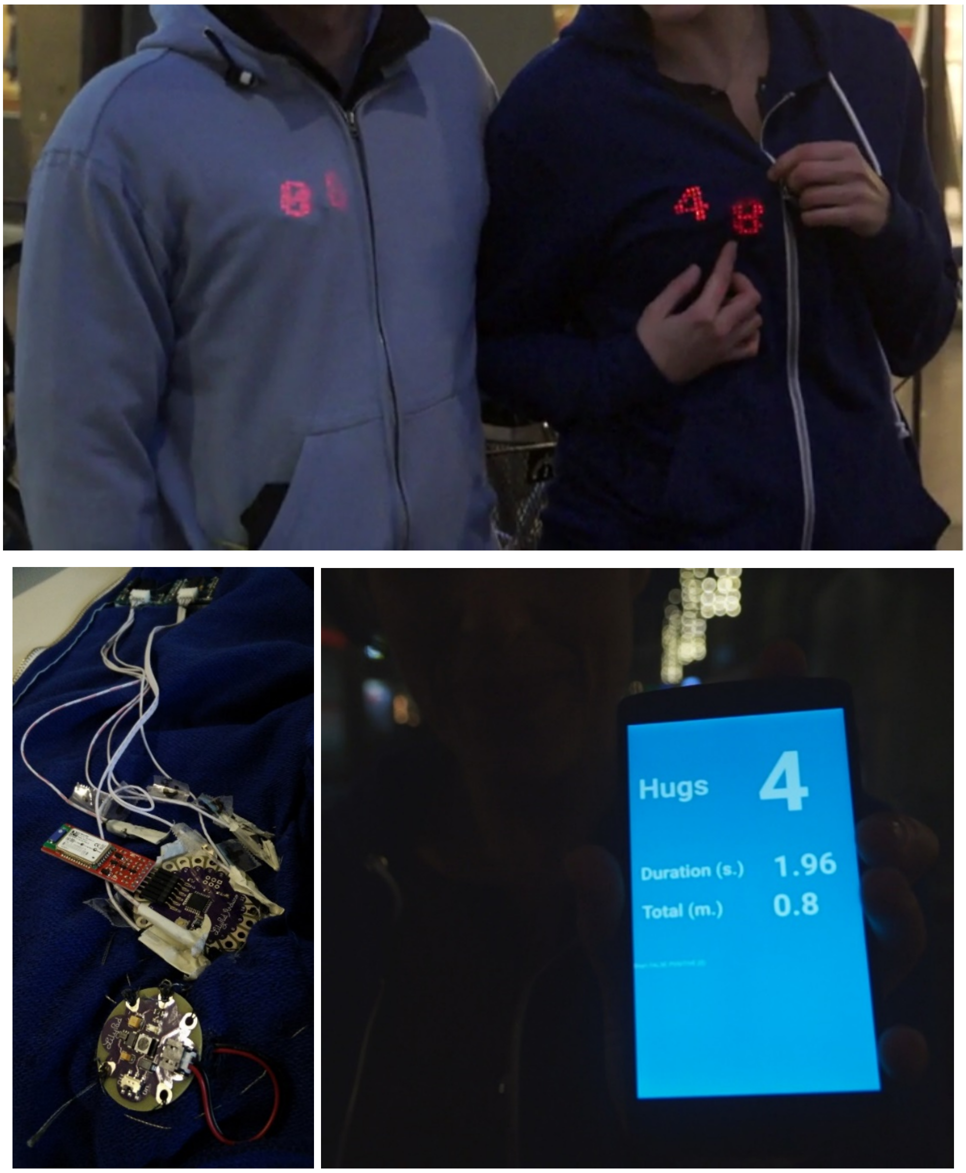

Figure 56. Hugginess hoodie prototypes and app

7.4.6 Evaluation

The first Hugginess system, i.e., the t-shirt, has been tested in a laboratory setting to assess, as a first step, the feasibility of the proof of concept. The system is able to recognize correctly most of the hugs, at least those executed with enough amount of pressure. The exchange of messages while hugging works almost flawlessly when the t-shirts are connected to the same ground, for example when powered through the USB connection of the same PC. Unfortunately, when the t-shirts are powered by the batteries, there is a lack of a common ground reference and the serial communication between the two conductive patches is no more reliable. To circumvent this problem, another conductive patch can be added to the t-shirt to share the ground between the two huggers. Supposing that the heads of the huggers are at the same height, this patch could be positioned horizontally on the top of the other two patches vertical patches for TX and RX. However, this solution adds complexity to the design and to the routing of conductive threads. Therefore, it constitutes an additional variable for a system that is already subject to several sources of electric noise. A preliminary test showed that adding a third patch solves the issue of the missing common ground, but, unfortunately, no further integration has been carried on for this system. Since the exchange of messages through the fabrics was the most complex part of the t-shirt, the second system (the hoodie) has been developed without this feature and recognized gestures only with pressure. This allowed obtaining a much more reliable prototype that has been used in several demonstrations. This prototype was also much more practical to test and showcase, since spectators do not need to wear their own t-shirt. A formal evaluation during the Free-Hug event in Bern in December 2015 had been planned, with the purpose of asking to people passing-by to test the system and share their feedback about this concept. Unfortunately, for logistic reasons, it was possible to reach the event location only when the event was almost finished, making hard to find people that would stop for a hug in the middle of a street, especially when the organizers of the event with the free-hug advertisements left the street.

7.5 Smart Watch[29]

7.5.1 Introduction and Field Requirements

The aim of this project is exploring tangible gesture interaction with everyday objects in the context of a smart home. As discussed in Section 3.2, these everyday objects can serve as tangible cues to digital

[29] This work has been published in 2012 in [142], in collaboration with Fondazione Bruno Kessler (FBK), in Trento (Italy).

content [146] or to controllable home functionalities [195]. In the previous Memodules project [146], the objects are used as simple token for digital content, obliging the user to put the object on a control board, the aim of this project is introducing the expressivity of gestures to interact with these objects as well as allowing the user to interact on the move, without the need of reaching an external device. This restrained the design space to a wearable device, which should be able to recognize objects and to recognize user's gestures performed in relation to this object. Inspired by the work of Feldman et al. [63] and Stikic et al. [196], a wrist worn device, in the form of a smart watch has been selected as a concept to explore for the recognition of tangible gestures with everyday objects.

A requirement of the system was to recognize different tagged objects and tangible gestures that users could perform with these objects to navigate media content associated to them. Moreover, the interaction with these objects should be economical, i.e., not fatiguing for the user.

7.5.1 Explored TGIF Aspects

This system allowed exploring the *control* application domain, specifically, for controlling the multimedia system of a smart home. The interaction was designed using the *expert-driven* method, with a *reification* of object (picking existing objects at home) and gestures and the application of the *polymorphism* principle to reduce the number of total gestures. In particular, for the design of this system the users where implicated before the implementation of the system (with a Wizard of Oz experience) to compare two different tangible gesture set, with or without the *hold* component. Since different objects where chosen, the preliminary study allowed also to investigate the impact of the *form dimension* and *deformability* affordances provided by the different objects on the ergonomics and user experience of the different gestures, as well as *semantic* affordances that could help to associate meaning to objects through a *noun* metaphor. The system has been designed according to the *wearable* approach and different technologies have been explored to recognize different type of tangible gestures. The design space of TGIF explored through this system is represented in Figure A8 in Appendix A.

7.5.2 Design Rationale and Methods

The design of the smart watch system followed an expert-driven approach, but with early assessments of the interactions, before the development of the system. Since the project aimed at gesturing with objects associated to digital media content to control a media-player application, a wearable solution (a smart watch) able to recognize tangible gestures with tagged objects was conceived. As discussed in Section 6.1.2, a practical wearable solution for recognizing everyday objects is attaching a small RFID tag in the object and reading it with wrist-worn RFID reader, while most forearm gestures can be detected through an accelerometer integrated in the wrist-worn device. This led to a first study to understand user preferences regarding gesturing with objects. Indeed, while in Feldman et al. research [63] objects were only used to select items in an audio based menu, without holding them while interacting with the media content, I was interested in evaluating the possibility to perform gestures while holding the tagged object in the hand (hold+move gestures). Therefore, before developing the system, I investigated the user appreciation of gesturing with physical artifacts through a Wizard of Oz experience. Eight participants (students aged 20 to 26, 3 female, all right-handed) interacted with four different objects linked to digital media. I asked users to wear a fake prototype of the smart watch (Figure 57, right) and to interact with a media player performing gestures with the objects. These gestures were designed to be recognized easily through an accelerometer and the syntax and semantics of these tangible gestures are discussed in the next subsection. The aim of the study was evaluating the user satisfaction with the systems, understanding if gesturing with objects in the hand could be cumbersome for the users (in a 5-point Likert scale) and investigating the appreciation of interacting through personalized gestures. After an analysis of the results of the experience and, in particular, after observing that all users would appreciate interacting through personalized gestures and that two users, without being instructed, performed a squeeze (which was not defined in the gesture set) for the gesture associated to quitting the media application, I decided that it was worth exploring a larger variety of gesture with objects in respect to those proposed for the Wizard of Oz experience. Since recognizing forearm gestures (in this context, hold+move gestures) with a wrist-worn accelerometer was already broadly explored in literature [63], [167], [3] and well validated in terms of reliability, I decided to investigate the possibility to recognize a different and more challenging set of gestures, those that do not imply a movement of the forearm. This set includes gestures performed with fingers while holding the object in the hand (hold+touch+move), different way to grasps an object (hold+touch) and gestures that apply forces to the held object, such as squeezing (hold+ pressure attribute). Inspired by the pioneer work of Rekimoto [170], I decided to investigate the recognition of gestures with objects through the analysis of the shape of the wrist. This topic has already been investigated with different technologies, such as capacitive sensors [47] or infrared sensors [130]. In this project, in collaboration with the Fondazione Bruno Kessler of Trento, I explored the possibility to use pressure sensors to detect changes in the shape of the wrist, associated to the aforementioned set of gestures. The implementation details of this project are shown in Section 7.5.4, while the results obtained with this system and with the Wizard of Oz experience are discussed in Section 7.5.5.

Figure 57. The four objects (left) and the mockup of smart watch used in the Wizard of Oz experience

7.5.3 Interaction Design: Tangible Gesture Syntax and Semantics, Feedback

While in the systems presented in Section 7.2, 7.3 and 7.4, the user can perform gestures with only one object, in this system, the users can interact with several objects. As in the Memodules system [146], each object can serve as a tangible cue to digital media or to a system function, thus its physical form can serve as a reminder and has a direct semantic reference with the digital content (*noun* metaphor). For the Wizard of Oz experience, in order to assess the users' appreciation of gesturing with objects in the hand, I have chosen four objects that could serve as tangible cues for particular media: a seashell for videos from Florida, a snowball for photos from the winter, a plush owl for photos of animals and a CD with classical music (Figure 57 left). For each object, the users could perform six tangible gestures; each gesture was associated to a command of the media player. To play the content of the object, users had to perform a shake gesture towards the TV. Users could browse the media list through a swipe of the hand to the left or to the right and they could raise or decrease the volume swiping up or down. Raising the hand and maintaining it in a vertical position was associated to the stop command. Besides these six gestures, obviously the user had to hold the object to select the desired media to which it was associated. During the experience, the users had the possibility to perform gestures holding the object in the hand (hold+move gestures) or without the object in the hand (move gestures) after selecting the object (hold or touch gestures).

Visual and audio feedback was provided in a TV in front of the user, displaying media according to the tangible gestures performed to the user.

7.5.4 Implementation

As anticipated in Section 7.5.1, a smart watch for the recognition of tangible gestures should deal with three main tasks: the recognition of objects, the recognition of movements of the forearm (move, hold+move), the recognition of grasps and of grasp with applied pressure, such as for the squeeze gesture (hold+touch&pressure attribute), and of other gestures that can be performed without moving the forearm, such as finger gestures (hold+touch+move) wrist extension and flexion (hold+move). Based on the following analyses, in [142], I proposed a diagram block for a smart watch able to deal with the aforementioned aspects of tangible gesture recognition. The diagram is shown in Figure 58.

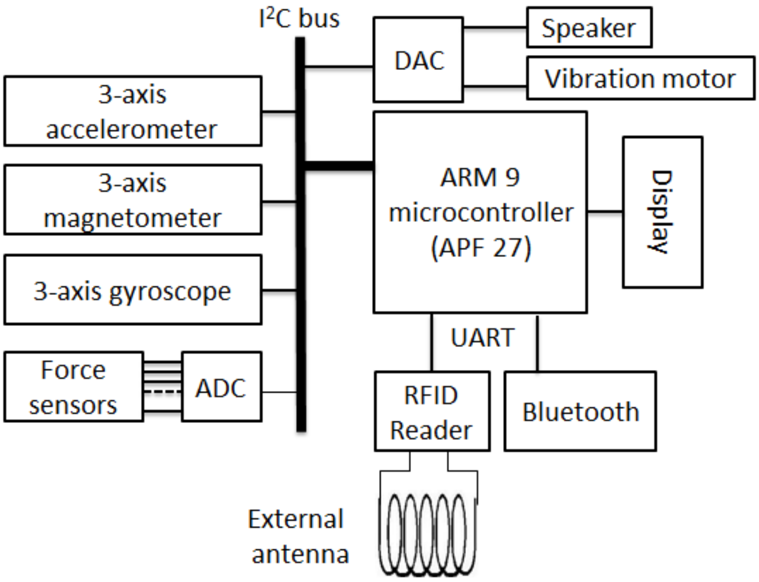

Figure 58. Block diagram of the proposed smart watch for recognizing tangible gestures.

The proposed smart watch is equipped with sensors both in the strap and in the body. A large strap is needed to enclose the antenna of an RFID reader. In the inside part of the strap, an array of pressure sensors allows to assess the wrist shape and to detect tendon movements. In the watch body, under the display, an ARM9 microcontroller running the Armadeus Linux embedded system elaborates data coming from sensors. The initial choice of this platform has been discussed in Section 6.2. The watch body should embed also a tri-axis accelerometer, magnetometer and gyroscope, positioned in the center of the watch; an Analog-to-Digital Converter for pressure sensors; an RFID reader; a vibration motor for haptic feedback and a speaker for audio feedback. A Bluetooth module allows the transmission of either raw data or gestures recognized from the embedded microcontroller.

As discussed in Section 7.5.1, objects that are held in the hand can be recognized by attaching a passive RFID tag to the object and integrating an RFID reader in the smart watch. The biggest challenge for this task is embedding an antenna with sufficient gain in the structure of the watch, more specifically in the belt, in order to recognize tagged objects held in the hand. RFID readers for passive tags generally work at three standard frequency bands:

- Low Frequency (125 KHz)

- High Frequency (13.56 MHz)
- Ultra High Frequency (433 MHz – 2.4 GHz)

In the literature most projects used LF [183] or HF [63, 67] readers and relative tags. HF tags generally allow achieving higher distances with less power, but they suffer from interferences, especially if metallic objects are present. During my thesis, I focused my investigations on two commercial development modules, the 125 KHz ID Innovations ID-2 reader and the 13.56 MHz SkyeTek M1-mini module, both with a nominal consumption of about 15 mA at 5 V. Having a lower price and being more robust to imperfections of the antenna, I chose the ID-2 module for rapid and economic prototyping of a 125 KHz RFID reader with external antenna. The antenna should have an inductance of 1.08 mH or less. With an inductance of 1.08 mH, the antenna will match the RFID reader internal capacitance of 1500 pF, giving a theoretical resonance frequency of 125 KHz, according to the following equation. In case of a lower inductance, the resulting frequency can be obtained increasing the value of the capacitance by adding an external capacitor.

$$\boldsymbol{f} = \frac{1}{2\pi\sqrt{LC}} = \frac{1}{2\pi\sqrt{1.08*10^{-3}*1500*10^{-12}}} = 1.25*10^{5}$$

In order to improve the reading range of the RFID reader in the direction of the forearm, I chose to use a solenoid antenna, which should be integrated in a large strap for the watch. The inductance of the solenoid is given by the following formula.

$$L = \frac{d^2 * n^2}{18d + 40l}$$

where L is inductance in µH, d is coil diameter in inches, l is coil length in inches, and n is number of turns. Inverting the equation and fixing a length of 10 cm (3.94 inches) and a radius of 4 cm (1.57 inches), we get that 153 turns are necessaries to obtain an inductance of 1.08 mH (1080 µH).

Different tags should be tested in order to obtain an appropriate range. Generally, bigger tags, like RFID cards, allow obtaining longer ranges, but obviously are difficult to attach to small objects. Moreover, the orientation of the tag is very important and should be perpendicular to the axis of the solenoid. Preliminary tests have been conducted with the ID-2 reader and a commercial rectangular antenna with dimensions of 8.5x12cm, and with a custom-built solenoid antenna based on the previous calculation (the results are presented in Section 7.5.5). The custom-built antenna is shown in Figure 59.

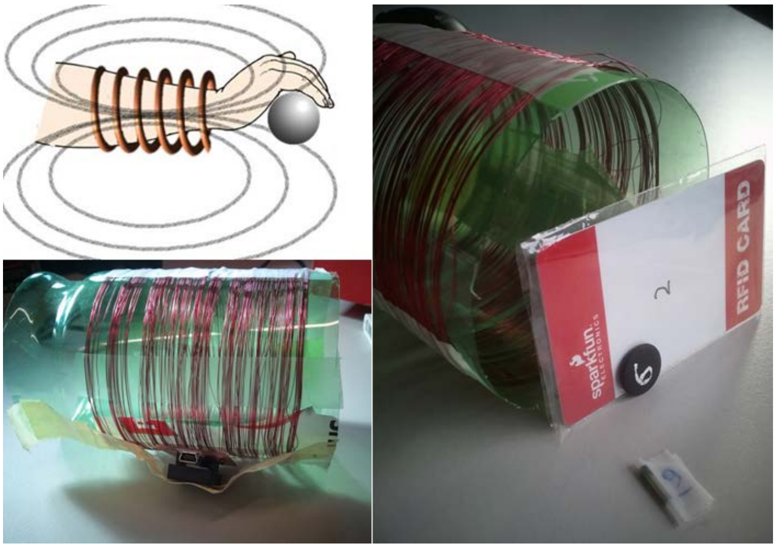

Figure 59. The custom-built antenna for reading RFID tags attached to objects in the hand, on the right, the antenna compared to popular RFID tags.

My investigations for the recognition of forearm gestures were limited to the setup of inertial sensors in the aforementioned ARM9 platform: the chosen inertial sensors were the 3-axis ADXL345 accelerometer, the 3-axis HMC5843 magnetometer and the 3-axis ITG-3200 gyroscope. These three digital sensors were connected through the I2C bus to the ARM9 microcontroller. As mentioned before, a particular challenge of this project was the recognition of all gestures that could not be detected through inertial sensors. To this purpose, flexible force sensors have been chosen to detect the movements of the tendons. As proof of concept, the sensors have been purchased from Tekscan [FlexiForce® A201]. The working principle is based on resistance changes due to the applied force on a piezoresistive ink between two electrodes on a flexible substrate. The diameter of the active area of the sensors is 1 cm. A piece of silicone has been glued on the sensible area of the sensor for a better transduction of the force exerted by the tendons.

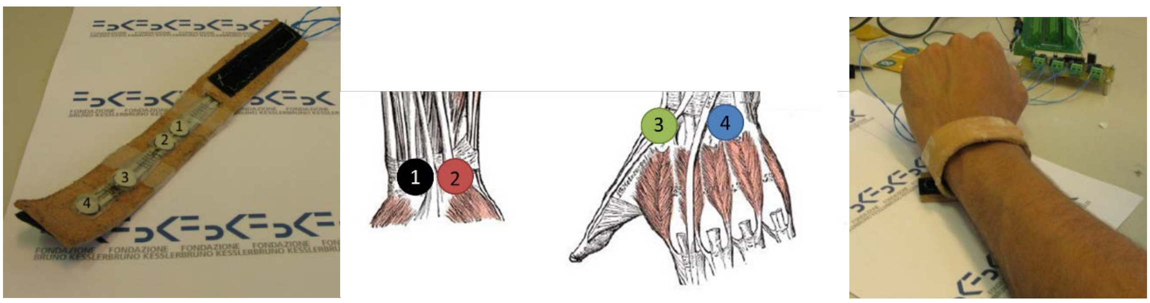

Figure 60. On the left, the wrist strap with four FlexiForce sensors mounted on it; in the center, the placement of the sensors in relation to sensors; on the right, the strap wore on the wrist and the conditioning electronics.

For the experimental purposes, four sensors have been mounted on a leather strip, granting the possibility to move sensors when needed, in order to fit different wrists (Figure 60). The position of the sensors has been chosen according to the forearm anatomy: sensor 1 is close to the tendon of the flexor digitorum superficialis tendons muscle, or next to the palmaris longus tendon, when present; sensor 2 is on the flexor carpi radialis tendon; sensor 3 on the extensor pollicis brevis and sensor 4 on the extensor digitorum. The sensors can detect forces up to 1 lb. with a sensitivity that depends upon the polarization circuit. The control electronics consists of a voltage divider with a rail-to-rail operational amplifier (LTC2201) and variable gain adjustable through a trimmer (Rg). The output voltage (Vout) is then converted in digital signal using an I2C-based Analog to Digital Converter (PCF8589) and sent to the ARM9 microcontroller for further elaboration. For the preliminary tests, analog signals have been acquired using an acquisition board (National Instrument NI PCI-MIO-16E).

7.5.5 Evaluation

In this sub-section, I present the results of the investigations conducted for the smart watch for tangible gesture interaction. First, I will present the results of the Wizard of Oz experience, then, the results of the preliminary investigations for object recognition and for the detection of tendon movements from the wrist. In order to assess the user appreciation of gesture interactions with objects, the eight participants to the Wizard of Oz experience had the opportunity to interact with a media player through the four objects and the tangible gesture set described in Section 7.5.4. Each participant completed a set of 30 predefined tasks that allowed exploring all the six tangible gestures for browsing the digital content associated to the four objects. In particular, each user completed two sets of 15 tasks: in the first set gestures were performed holding the object in the hand while gesturing (participants were instructed that tangible gestures were recognized through the fake smart watch); in the second set gestures were performed selecting first the object to interact with (by holding it or touching it) and then performing the gestures without holding the object (participants were instructed that in this case gestures were recognized by a camera). After the two sessions with the two different systems, I asked users three specific questions about the media content, which forced the users to find the answer in the media. In this case, I informed the users that they could use their preferred system, thus they were free to perform gestures either with or without the object in the hand. I noted the choice done for answering each question and for performing each gesture. After the experience, the participants rated the following sentences in a 5-point Likert scale (1 strongly disagree – 5 strongly agree):

S1. Gestures with objects in the hand are comfortable
S2. Gestures with objects in the hand did not imply a physical effort for the user
S3. Free-hand gestures are comfortable
S4. Free-hand gestures did not imply a physical effort for the user

S5. Gestures with objects in the hand imply a cognitive effort to remember or execute them

S6. Free hand Gestures imply a cognitive effort to remember or execute them

S7. Performing gestures with your non-preferred hand is cumbersome

S8. Performing gestures with your secondary hand could be useful (for example to execute other tasks at the same time)

S9. A set of gestures defined by the user could be useful

Answers to S1-S9 are presented in Figure 61. All the 8 participants wore the watch on their preferred hand. Although they generally think that performing gestures with the non-preferred hand is not very comfortable (S7), they would like to have the possibility to choose the hand where the watch is worn. From (S1-S4) it can be seen that gesturing without objects is preferred, with an average rating of 4.75 (SD 0.46) and 4.37 (SD 0.74) versus only 3.00 (SD 0.75) and 3.37 (SD 0.92). In particular, some users reported pain in gesturing with objects. One user found the snowball too heavy to perform several gestures and another affirmed that the CD was less comfortable to hold. Some participants found that gesturing with objects implied less cognitive effort to remember gestures (S5-S6).

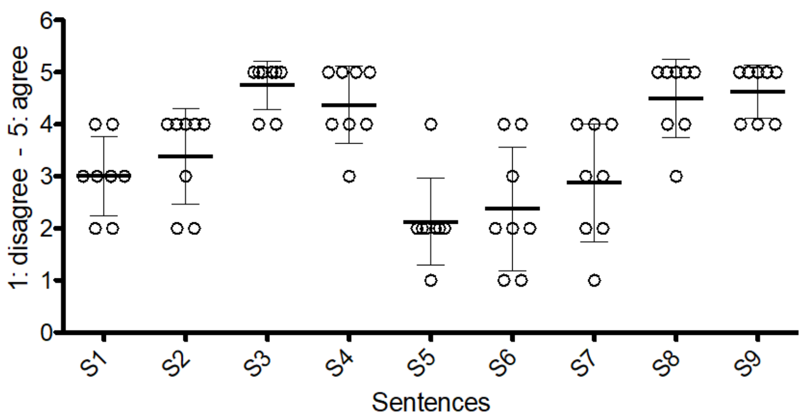

Figure 61. Results of the questionnaire for the Wizard of Oz evaluation.

In the third part of the evaluation, when users could choose whether gesturing with objects in the hand or not, 5 people always left aside the objects, 1 person held the object for all gestures, 1 only for the "play" gesture and 1 only for the second and third question. All people agreed that defining their own set of gestures would be useful although someone affirmed that the watch should have a set of predefined gestures for basic tasks. Many gestures suggested by the participants are related to free-hand gestures (like pointing to an object or a household device), or gestures to control a Windows environment. They also suggested whole body gestures like dancing or yawning to select appropriate music. For the snowball, someone suggested looking into the ball to start playing. Another participant suggested to open/close the CD to play/stop music.

Although most users found gesturing with objects more cumbersome than free-hand gestures, some of them chose to keep the object in the hand during the third phase of the evaluation. I believe that in many cases it could be a natural way to interact if proper gesture-object pairs are chosen. In fact, the gestures proposed for the user evaluation, rather than being related to the objects, were generally coupled with the output in the screen, limiting the purpose of the object to only the media selection function (cf. functional gestures of Carrino et al. [42]). Following these considerations, it is worth noting that some users proposed more meaningful gestures that better exploit the object affordance (such opening/closing the CD, looking into the snowball, etc.) obviously, while such gestures are probably easier to remember and more intuitive for the users, they cannot be associated to objects with different affordances, therefore enlarging the number of gesture-object pairs that the user should remember to operate the system. As suggested by some users in [123], probably the best option would be to define a predefined gesture set for basic operations in the system and then allow the user to define their own gestures for particular objects or functions.

Concerning the object recognition with a wrist-worn RFID reader, the first test conducted with an ID-2 reader and the commercial rectangular antenna allowed recognizing only card tags at a range of about 4 cm, while the other smaller tags could not be recognized. With the aforementioned custom-built solenoid antenna, the detection range of card tags increased up to 9cm. Moreover, with this antenna the ID-2 reader managed to detect smaller tags, obviously at a lower range. Finally, using another sample of the ID-2 reader, I obtained similar results also with the commercial antenna, showing that the antenna should be fine-tuned for each sample in order to maximize the reading range. As expected, the detection range decreased noticeably according to the inclination of the tags: the maximum range can be obtained when the axes of the reader antenna and of the tag antennas are parallel.

Preliminary tests have been performed in order to check the possibility of discriminating different positions and gestures by means of force sensors. Several gestures have been considered and the results are reported in the Figures 62 to 64.

First, free hand gestures, such as flexion and extension, radial deviation and ulnar deviation, pronation and supination were analyzed, as shown in Figure 62. The radar representation in Figure 62 (right) shows that the recognition of the movements is possible by comparing the signals obtained from the four sensors. The radar graph was obtained using a normalized mean value.

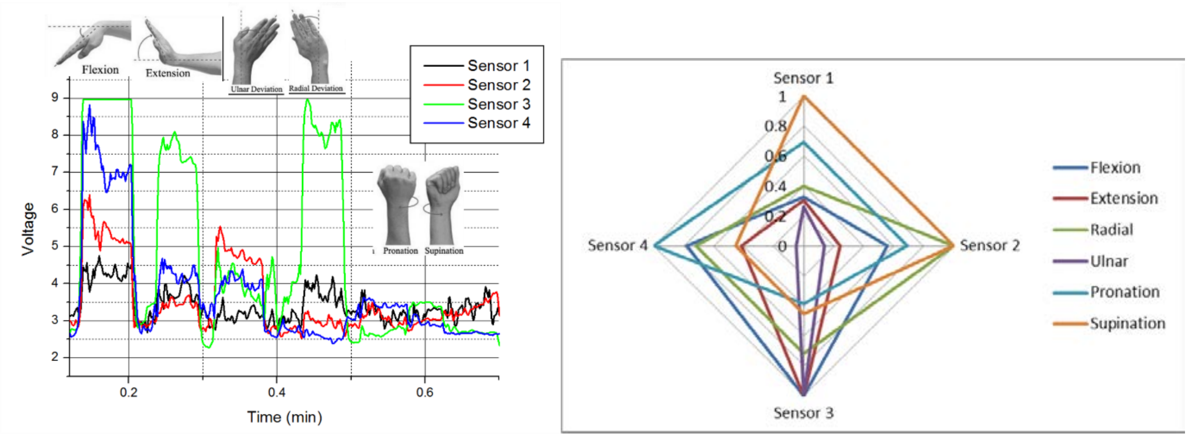

Figure 62. Free hand gestures. Acquired signal from the sensors in the different positions (left) and radar representation of the normalized mean values (right).

It should be noticed that sensor 3 on the *extensor pollicis brevis* tendon has great variations while flexing, extending and in the radial deviation due to the movement of the thumb. Other relevant contributions are from the *flexor carpi radialis* tendon (sensor 2).

The second set of gestures aims at investigating the possibility to recognize the hold gesture and the squeeze gesture. To this purpose, the four pressure sensors have been tested through the following actions: closing the hand, opening the hand, squeezing a spherical object and relaxing the hand (Figure 63). In this case, the main variations are related to the activity of the *extensor* (sensor 4) and *flexor* (sensor 1) *digitorum* tendons in the open-close sequence, while sensor 3 is important in discriminating the squeezing gesture from relax position.

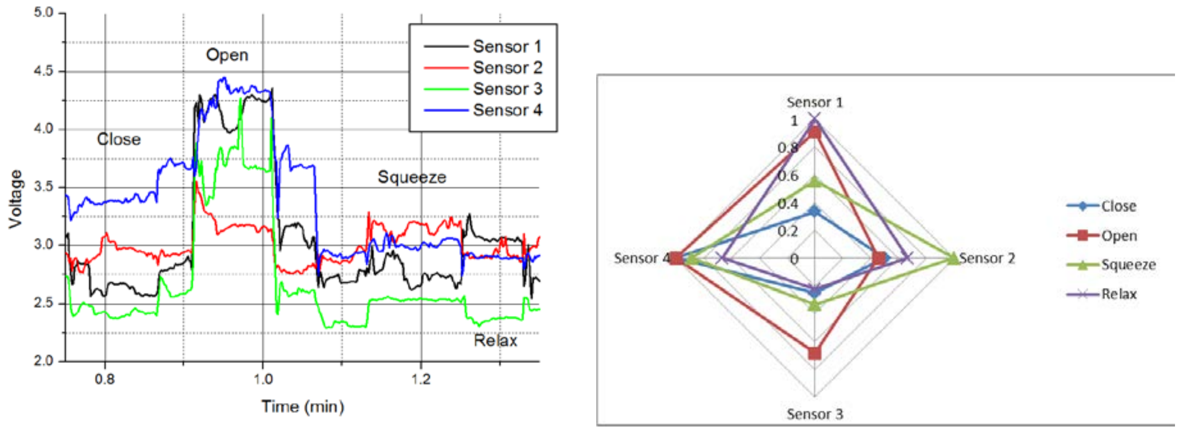

Figure 63. Segmenting gestures experiment. (a) Acquired signal from the sensors in the different positions and (b) radar representation of the normalized mean values.

Finally, the grasp of a pen has been studied as a particular case of the recognition of different hold+touch gestures; it is possible to distinguish between tip grasp and spherical grasp [179] (Figure 64). In the graph in Figure 64, the peak between the positions appears during the transition, while changing the pen grasp.

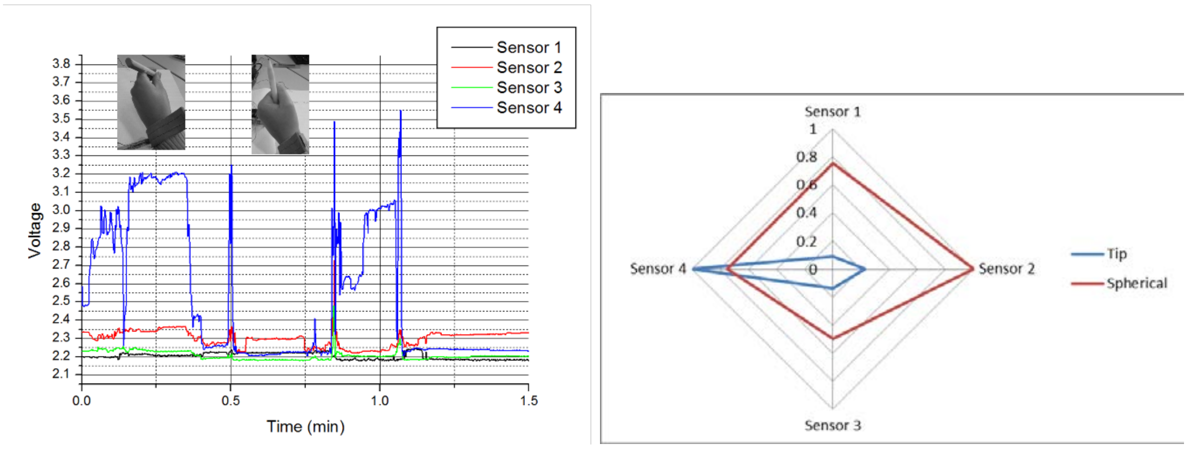

Figure 64. Tip grasp and spherical grasp. Acquired signal from the sensors in the different positions (left) and radar representation of the normalized mean values (right).

These preliminary tests show that the four force sensors are able to detect wrist and hand postures. However, the actual shape of the signals strictly depends upon several factors, which introduce a level of incertitude in the measurement. In fact, it is difficult to place the sensors always in the same location with respect to tendons and to reproduce the same position of the hand and of the wrist. Each imperceptible deviation of the hand position has a considerable impact on the response; moreover, wrist shapes and muscle configurations can vary a lot from one person to the other, affecting the recognition procedure. This latter problem could be avoided by training gesture recognizers only on the user that will use the system, which is an acceptable approach for a personal device like the smart watch. It should be noticed that the used sensors are quite big compared to the tendons and they can cover only a few positions around the wrist. Therefore, further implementations with a higher number of smaller sensors could lead to better results.

In the end, this project remained merely exploratory, since it has not been possible to build a functioning smart watch, because of the obvious technical difficulties. Meanwhile, in the last years, many smartwatches have been released in the market and I investigated the possibility to use an existing smart watch to recognize gestures with objects. Unfortunately, among all the smartwatches, only the Sony Smartwatch 3 implemented NFC technology (similar to RFID technology), which could be used to recognize tagged objects. Unfortunately, there was no API available to use the smartwatch as a NFC reader, since NFC was used only for pairing with the smartphone[30].

7.6 Other Related Systems and Activities

During my thesis, I supervised and/or contributed to the design of several other tangible interactive systems. Unlike the other systems presented before, I will only briefly describe one of these systems in this section, as a further evidence of the investigations conducted in the field of TGI. Moreover, two workshops have been conducted in a topic strictly related to this thesis and four other workshops have been conducted in a closely related domain. This work will be also presented in this section.

With the purpose of exploring the application domain of digital content production within the work environment, during my thesis I worked also on the design and development of a smart pen. The purpose of this project was to investigate the usage of a smart pen as a multipurpose tool that changes function according to the way it is grasped (*hold+touch*). The aim of the system was to increase the productivity in work environment and the interaction design exploited the *polymorphism* design principle to change the *identity* of the pen according to the gesture that is performed on it. An example of the possible hold+touch gestures and the possible meanings obtained through the *verb* semantic constructs are shown in Figure 65. The design space of TGIF explored through this system is represented in Figure A9 in Appendix A.

Figure 65. Eight different pen grasps (left) and the corresponding verb metaphors (right)

[30] A custom firmware to use the Smartwatch 3 as a NFC reader has been recently released by a member of the XDA developers forum and, although it has many limitations, it could be probably used to recognize NFC tagged objects.

This original idea, conceived at the beginning of my thesis, was implemented and published few months later by Microsoft Research in 2011 [192]. Nevertheless, some investigations have still been conducted: following an *embodied and embedded* approach, I investigated the possibility to implement a capacitive grid to cover the pen and recognize grips and gestures; in a further study, a simpler smart pen based on 12 circular electrodes connected to a MPR121 capacitive sensor has been developed. The pen was connected to an Armadeus ARM9 platform, which recognized 12 tripod grips at different height of the pen to change inking mode and one palm grip to use the smart pen as a rubber. A simple drawing application was developed for a tabletop interface. A video of the working system can be seen at the following address: https://youtu.be/XqAprwWV_tw

As additional activities related to my thesis, it is worth mentioning two workshops that have been conducted on the topic of tangible gesture interaction. The first workshop was held during the French German Interaction Studio in 2013 in Bidart (France). The aim of this workshop was to test the abstracting part of the TGIF framework and in particular the gesture syntax with young researchers and practitioners working in the field of tangible interaction. Participants were asked to express their preference according to the different tangible gestures (hold+move, Hold+touch, touch+move) to be used to interact with a lamp. Participants were divided in three groups and each group designed developed an interactive lamp according to the chosen type of tangible gesture. The lamps were developed with Arduino and different sensors to recognize the different gestures. Examples of the developed prototypes are shown in Figure 66. The experience was generally positive and participants enjoyed developing the different prototypes with Arduino. Some difficulties in the comprehension of the gesture taxonomy and of the different gesture types were manifested by participants, which helped me to improve the way the taxonomy should be presented.

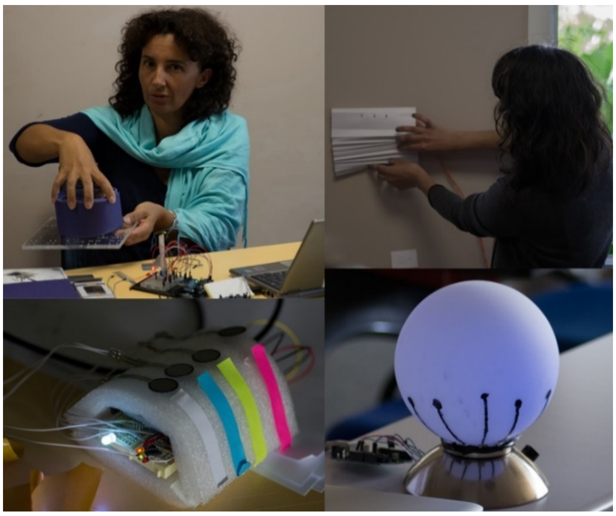

Figure 66. Prototypes realized during the FGTIS workshop.

The second workshop was held during the Tangible Embedded and Embodied Interaction (TEI) conference in 2015 in Stanford (CA, USA). The workshop was organized in collaboration with Elise van den Hoven and Ali Mazalek, the authors of the pioneer article on TGI [213]. The aim of this workshop was discussing Tangible Gesture Interaction in a larger perspective, in particular on the possibility of blending tangible interaction and gesture interaction into a new paradigm. During an explorative ideation session, the participants were divided in three groups and designed three tangible gesture interactive systems in three particular application domains (interaction with plants, interaction with pets and interaction with aliens). The results were very interesting in terms of the designed interactions, thanks also to the singularity of the three application domains. The mockups of the prototypes are shown in Figure 67. Moreover, four position papers were presented during the workshop and I also presented a position paper on the abstracting part of TGIF. The discussion with workshop participants was very useful to individuate some peculiarities of the framework, such as the user-centered aspect of the taxonomy and the difficulty to describe object-centered gestures (this limitation is discussed in Section 8.2.2). As a follow-up to the workshop, a special issue on Tangible Gesture Interaction in the MDPI Machine journal has been edited by the organizer of the workshop. More details about this workshop and the special issue can be found at this address: http://www.tangiblegestures.com/

Figure 67. Mockup realized during the Tangible Meets Gestural workshop at TEI'15

Finally, it is worth mentioning also the organization of three workshops (at Ubicomp 2015, TEI 2016 and Ubicomp 2016) on the topic of Full-Body and Multisensory interaction, as a natural continuation of my investigations on the design of richer interactions that brings back the human to the real world and that better exploit human innate skills. This series of workshop allowed to explore richer gesture interactions with the real world, that take advantage of the full body, as well as the exploitation of non-conventional senses for providing richer feedback and for exploiting the peripheral attention of the user. Additional information about this series of workshops can be found at this address: https://sites.google.com/site/bodysenseux/

My latest investigations concern the exploration of tangible interaction as a richer interaction means for the Internet of Things (IoT). While most IoT objects can be accessed only remotely with smartphone or web interface, there is the evident need to bring back the interaction with these object in the real world, having richer in-situ interactions that better exploits the innate human skills. This topic has been discussed during the second European Tangible Interaction Studio, which I co-organized in Fribourg. As a follow-up to this activity, I co-organized the workshop on Tangible Interaction with Light in the IoT, at Ubicomp 2016. More information on this workshop can be found at the following address: https://sites.google.com/site/tangiblelighting/.

7.7 Conclusion

In this chapter, I presented seven different tangible gesture interactive systems for three application domains. Each system investigated different aspects of the Tangible Gesture Interactive Framework, advancing at the same time the existing research in the specific application domain.

In particular, for the interaction in the car, four systems have been developed, investigating three technological approaches (wearable, embedded and embodied, hybrid) and two different methods for the interaction design (expert-driven, according to technological constraints, and user-driven, thanks to a gesture elicitation study). The originality of the approach of gesture interaction on the steering wheel has granted different publications in the AutomotiveUI conference and a remarkable interest also from the automotive industry. Moreover, this original interaction modality has been compared with two existing interfaces for interacting in modern cars: vocal interaction and touch interaction on the central dashboard.

For the domain of interpersonal and emotional communication, two systems have been developed, investigating two different technological approaches (embedded and embodied, wearable). In particular, the ADA lamp design was based on best practices for life-like objects and anthropomorphic affordances as well as on the results of previous studies that I conducted in this domain. Conversely, the Hugginess system highlighted the importance of physical touch in interpersonal relations, augmenting and rewarding the hug gesture during physical encounters. It is worth noting that the ADA lamp system won the best Work-In-Progress prize at the IHM'14 conference and the Hugginess project received media coverage by the Swiss television.[31]

Finally, for the domain of the interaction in the smart home, an additional wearable system, i.e., a smart watch, has been investigated for the interaction with everyday objects. In this context, a Wizard of Oz experience has been conducted in order to investigate with users the design space of tangible gesture interaction with everyday objects associated to digital content. Furthermore, two feasibility studies have been conducted in order to assess the possibility to recognize tagged objects with a smart watch and to recognize particular hand gestures and grasp postures detecting tendon movements with pressure sensors integrated in the watch strap.

Analyzing the different systems in terms of interaction design, I could notice that according to the application domain the requirements on the type of tangible gestures that should be designed may vary considerably. Indeed, while in the control application domain (WheelSense and Smart Watch) users may desire simpler gestures, easier to learn and to perform and often linked to legacy interactions they already know from other systems, in the affective communication domain (ADA Lamp and Hugginess) richer tangible gestures are well appreciated, especially if those gestures aims at stimulating reflection and at soliciting emotional responses. This difference is reflected by the Heideggerian dichotomy (cf. Section 2.2.2.2) typical of tangible interfaces, constituted by interactions that are ready-to-hand, i.e., transparent to the user's mind and often executed in the periphery of attention, and by other interactions that are present-at-hand, which make the users reflect about what they are doing and the impact that their actions have on the system and more in general on the surrounding context, including other people.

Concerning the building phase, I could notice that developing a TGI system may require many different skills. In particular, the embedded and embodied approach and the wearable approach require electronic and manufacturing skills to embed the technology on the object or in a wearable device, in an efficient and well-designed package where technology disappears to the users' sight. At the same time, the recognition of gestures requires often good programming skills, especially when machine learning is required to classify complex gestures. Toolkits, prototyping platforms and gesture recognition frameworks can help speeding up the development of TGI systems, providing pieces of hardware and software that can be reused in different contexts. However, the heterogeneity of possible tangible gesture, the three different building approaches and the large variety of application domains make very difficult the realization of a generic plug-n-play toolkit for building whatever TGI system.

[31] http://www.rts.ch/play/tv/12h45/video/minimag-visite-au-salon-human-tech-a-fribourg-ou-les-besoins-de-lhomme-sont-au-centre-des-preoccupations?id=6469965

Nevertheless, even a toolkit that could allow recognizing a small set of tangible gestures through an embodied and embedded approach would be without doubt helpful for designing and developing TGI systems for everyday use.

8 Framework Evaluation

"La somma filicità sarà somma cagione della infelicità,
e la perfezion della sapienza cagion della stoltizia."
Leonardo da Vinci, "Aforismi, novelle e profezie".

This chapter presents an assessment of the Tangible Gesture Interaction Framework. After an analysis of existing criteria for framework assessment, three criteria proposed by Beaudouin-Lafon [23] have been chosen: the descriptive power, the evaluative power and the generative power. In particular, for the descriptive power, it presents a classification of tangible gestures from literature according to the TGIF syntax and semantics; for the evaluative power, it presents the analysis of the interaction qualities of two applications developed during this thesis; and for the generative power, it discusses the systems developed during this thesis according to TGIF.

8.1 Evaluation Methods

TGIF aims at guiding the design and development of tangible gesture interactive systems. Tangible gestures offer interesting opportunities in several application domains and are promising for a future where more and more everyday objects will become interactive. Although previous studies in the field of gestures with objects already exist, the TGIF framework offers a high-level overview on the large variety of tangible gestures and tangible gesture interactive systems that already exist or that can be designed. Demonstrating the usefulness of this framework is not easy without a broad adoption by researchers and practitioners of the field. Indeed, a proper evaluation would have required to propose the framework in its integrality to several designers of interactive systems. Unfortunately, it was not possible to conduct during my thesis this kind of evaluation: first, for a matter of time, since the framework has been completed only at the end of this thesis; second, for a lack of available designers that would have adopted this framework to develop a TGI system over several months. Actually, during my thesis, a master student in engineering did have the opportunity to design a TGI system following my framework. However, at the beginning of his project, the framework was composed only by the TGIF syntax, thus providing few guidance for designing and building TGI systems. The student managed to develop an interesting system (cf. the smart pen presented in Section 7.6) but he did not have much time to improve the designed interactions and to validate the system with users. Nevertheless, he reported that the TGIF syntax was clear and that it was helpful to understand the domain.

To the best of my knowledge, few frameworks found in literature have been validated through proper long-term testing with interaction designers. Among the most meaningful examples, in her PhD thesis, Soute [193] assessed her framework RaPIDO, a platform for rapid prototyping mobile outdoor games, through a longitudinal interpretive case study with two design master students working on the project for two months. It is worth noting that the framework was focused mainly on *building*, rather than on *abstracting*, therefore, it was easier to apply to a concrete project. Hornecker and Buur's framework on tangible interaction [95], which focused mostly on *designing*, has been validated presenting three case studies, two of them as a post-hoc evaluations of the system, and one where one of the authors participated to the design. In a later work, Hornecker proposed a set of cards based on the themes of their framework and validated it through additional 10 sessions of 45 to 90 minutes, 5 of them to evaluate already completed projects and 5 as designing sessions [93]. Wimmer's GRASP framework, which focused mostly on *abstracting*, assessed the formal validity of his framework according to the Vaandrager's definition of what is a *good model* [208]. Wimmer proposes also the analysis of two existing systems according to his model. While Vaandragers' definition deals with generic models, Beaudoin-Lafon [23] proposed a definition for a *good interaction model*. Since TGIF addresses specific questions concerning a new interaction model, i.e., tangible gesture interaction, for my thesis I have chosen this definition to assess the formal validity of the TGIF framework. According to Beaudoin-Lafon [23] a good interaction model should include three main dimensions: "(1) descriptive power: the ability to describe a significant range of existing interfaces; (2) evaluative power: the ability to help assess multiple design alternatives; and (3) generative power: the ability to help designers create new designs.". In the following sub-subsections, I propose an evaluation of TGIF according to these three dimensions. As many other framework authors did, I also presented in Chapter 7 several applications as case studies of my framework. While the details about how each system fits into TGIF have already been presented in the respective sections, their analysis according to TGIF is further discussed and summarized in Section 8.4.

8.2 Descriptive Power

For Beaudouin-Lafon [23], a good model should be able to describe a good range of existing interfaces. Indeed, one of the purpose of TGIF is to allow describing tangible gesture interaction according to different criteria. In particular, the abstracting section of the framework (Chapter 4) presented a taxonomy of gestures and analyzed them according to their syntax and semantics. In particular, in Section 4.3, I provided several examples that show how tangible gestures can be described through the three move, hold and touch components and additional attributes such as pressure, amount, speed, time, direction, *etc.* Section 4.4 presented a taxonomy of semantic constructs that can be used to associate meanings to tangible gestures. Moreover, Chapter 6 presented different approaches that can be used to describe the way TGI systems are built. The whole design space of TGIF is represented in Figure A1 in Appendix A.

As an additional and more complete validation of the descriptive power of TGIF, in Section 8.2.1, I present a classification of existing systems according to the TGIF dimensions and the result of a study conducted with 4 HCI experts to assess the inter-rate agreement of a classification of 13 tangible gestures according to the proposed taxonomy. I further discuss the validity and genericity of the descriptive power of the TGIF syntax in Section 8.2.2.

8.2.1 Classification of Existing Systems According to TGIF

Table 8 resumes the classification of tangible gestures found in existing systems. This classification does not aim to be complete—gestures with objects have been adopted in many systems—but rather, it offers several examples from all different classes depicted throughout the framework. For each system, I specified the component involved in gestures, *i.e.*, move (M), hold (H) and touch (T), the objects used for gesturing, the application aim, and the technological approach used to recognize gestures.

Table 8. Classification of TGI systems found in literature. EE = Embedded and Embodied, W = Wearable, E = Environmental. M= Move, H=hold, T= Touch.

System	Application	Tech Approach	Object (Affordance)	Gesture	Syntax M	Syntax H	Syntax T	Semantics
TabletopCars[52]	Gaming	E	Car toy	Move	x			Verb
Rosebud[75]	Storytelling	E	Doll or toy	Hold		x		Deictic
Video bubbles[178]	Storytelling	EE	Big bubble (deformable)	Hold&Pressure (Squeezing the bubble)		x		Embodied metaphor
Reading Glove[198]	Storytelling	W	Any object	Hold		x		Deictic
Pasquero *et al.*'s Watch[153]	Control	EE	Watch	Touch | touch + move			x	None
Picture this![217]	Storytelling	EE	Doll	Hold + move	x	x		Verb
Gesture Sticker[15]	Control	EE	Any object	Hold + move	x	x		None
Pen rolling[27]	Individual production	E	Pen	Hold+move (rolling the pen)	x	x		None
SplashController[73]	Gaming	EE	Different types of water containers	Hold+move	x	x		Embodied metaphor
MoSo Tangibles[19]	Individual production	EE	Various artifacts	Hold+move	x	x		Embodied metaphor
Graspables[199]	Control/Gaming	EE	"bar of soap"	Hold+touch		x	x	Verb
Ashbrook *et al.*'s watch[14]	Control	EE	Watch	Touch+move	x		x	None
TZee[226]	Collaborative production	E	TZee pyramid	Touch+move	x		x	None
Spinning in control[115]	Control	EE	Remote Controller (moving parts)	Hold + (Touch + move)	x		x	Verb
Hapticat[235]	Emotional design	EE	Cat (zoomorphic)	Hold + move | Touch + move | hold&Pressure	x	x	x	Noun+verb
MTPen[192]	Individual production	EE	Pen	Hold+Touch | Toch+move	x	x	x	Verb
Tickle[232]	Control	W	Any handheld device	Hold+(touch+move)	x	x	x	None
FoldMe[131]	Control	EE	Foldable display (deformable)	Hold+move (Hold+move)+touch	x	x	x	Noun+verb
dSensingNI[118]	Control/Collaborative production	E	Many objects in the environment	Touch | Touch+move | Hold+move	x	x	x	Deictic | Verb
PaperPhone[125]	Control	EE	Flexible Phone (deformable)	Hold+touch+move (bending)	x	x	x	Embodied metaphors

To better assess the validity of the proposed classification, during the definition of the move, hold and touch taxonomy, I conducted an evaluation with four HCI experts, who classified individually 13 gestures across seven classes: in this version, hold + move + touch gestures were divided in three subclasses and there was a class for non-TGIF gestures. I gave the experts a preliminary version of the TGIF syntax presented in Section 4.3, without any of the examples from the literature, which were part of the gestures to be evaluated. The description of gestures to be classified was extracted from the respective articles found in literature and reported without modifications. I have calculated the inter-annotator agreement rate of the independent evaluations using the free-marginal multirater kappa, obtaining a value of 0.63, which is already considered as a sufficient result. After a discussion among all the participants, the kappa coefficient of the agreement rate increased to 0.86. All HCI experts managed to understand the model and to classify the 13 gestures in about an hour, without prior knowledge of the model. Feedback received by the four HCI experts allowed to refine the move, hold and touch taxonomy, which has been simplified as shown in the current version of Section 4.3. Trading off the descriptive power of the framework with the simplicity of the description and easiness of understanding was a key challenge for the definition of the taxonomy.

8.2.2 Validity and limitation of the move, hold and touch taxonomy

Additional feedback on the taxonomy was gathered during the French German Tangible Interaction Studio, in 2013, and during the workshop Tangible Meets Gestural that I co-organized at the TEI'15 conference. From all these experiences, I was able to identify an important peculiarity of the proposed move, hold and touch taxonomy. Indeed, since the definition of the taxonomy has a user-centered perspective, gestures descriptions focus more on the users' bodily interaction with the objects rather than on the effects that such interactions have on the objects. For this reason, all those gestures that focus on the object movements, rather than in the interaction of the user with the object are more difficult to describe with the proposed taxonomy, as I conceived it originally. Indeed, in the proposed taxonomy, the move component is associated to a user's body part, which can be coupled or not to the object movement, depending on the integrality of the two movements (for example in many hold+move gestures). Some gestures that focus on the object movements, such as pen rolling [27], are difficult to describe in terms of hand manipulations, e.g., touching the pen in different points or moving the fingers in a particular manner in order to make the pen rotate, while holding the pen in the hand. For gestures that focus on object movements, a shift from a user-centered perspective to an object-centered perspective should be made. From a technological perspective, this shift of point of view to describe gestures (especially for the move component) facilitates also a shift for the building approach from wearable to embedded and embodied. Indeed, tracking object movements that are not directly associated to the movements of the hand is easier with inertial sensors embedded in the object than with other sensors integrated in a wearable device. Back to the pen rolling example, Bi et al. [27] exploited an environmental approach for recognizing this gesture. It is worth noting that also in this case, the system focused on the object movement, instead that on the hand manipulations performed by the user.

In conclusion, the touch, hold and move taxonomy can be also used to describe gestures that focus on object movements, although a shift of perspective should be adopted for these tangible gestures.

8.3 Evaluative Power

Evaluating different alternatives for tangible gesture interaction would be difficult without considering the specific application domain of the tangible gesture interactive system that the designer would like to build. Through the common practices for designing TGI (Section 5.4), I have shown how it is possible to choose among the broad range of tangible gestures according to the application domain or the affordances offered by everyday objects. Obviously, TGIF offers only an insight on existing common practices, but further explorations are needed on each application domain. In fact, the interaction designer of a new application should compare different gesture sets by conducting user evaluations. As discussed in Section 5.2 and as shown by Morris et al. for surface gestures [144], gesture elicitations are able to maximize the user's appreciation and the guessability of the proposed interactions. A comparison between expert-designed gestures and user-elicited gestures should be performed for the different application domains in order to ensure that Morris et al. results are still valid for different contexts, including different types of gestures.

In Section 5.3, I discussed how to evaluate the overall quality of the interaction of a tangible gesture interactive system, based mainly on the criteria of interaction qualities for tangible interaction, proposed by Hoven et al. [214]. As an example, in Section 8.3.1, I applied the aforementioned criteria to evaluate the interaction of the system presented in Section 7.3, i.e., the ADA lamp, and the systems presented in Section 7.2, i.e., WheelSense. It is worth mentioning that for a complete assessment of the system, beyond the interaction qualities, the designer should conduct several evaluations in order to assess system performances, system usability and other parameters that could be particularly relevant for the specific system, as discussed in Section 5.7 and in Section 6.1.4.

8.3.1 Interaction Qualities in the ADA lamp and WheelSense systems

In this subsection, I present an analysis of the designed interaction for the ADA lamp system (in the scenario in which the lamp behaves as a smart companion), and of the WheelSense systems, according to the six tangible interaction qualities described in Section 5.3. As discussed in Section 5.3, some interfaces might not respect, on purpose, one or more quality criteria. Indeed, the evaluation of the interaction qualities should serve as a critical analysis of the designed interaction, to individuate potential weakness or strong elements and possibly improve the design whenever needed, and not as a benchmark between different systems.

For the ADA lamp:

- *Integrated control*: The emotional state of the ADA lamp can be controlled performing gestures on the surface of the lamp. The sensors needed to recognize these gestures are also integrated inside the lamp.
- *Integrated representation*: The lamp emotional state is represented by its facial expressions, which are generated through RGB LEDs integrated in the lamp.
- *Direct control*: The lamp state cannot be controlled directly through tangible gestures. Indeed, although a deterministic state machine is used to describe the lamp behavior, I avoided obtaining direct reactions to user’s gestures. The almost unpredictable reactions are intended to create a life-like behavior that should foster long-lasting interactions.
- *Direct representation*: The representation of the lamp emotional state is direct, since there is only one representation for each emotional state. These representations are coded through specific facial expressions and colors.
- *Meaningful control*: The gestures used to control the lamp are those typically used to interact with humans. Therefore, the meaning and the emotional valence of gestures are consistent with the users’ social habits, at least in the context of the European culture.
- *Meaningful representation*: The representation of the lamp emotional state through facial expressions is meaningful and generally can be understood easily by the users. Obviously, the mapping between colors and emotions according to Plutchick’s wheel [161] of emotion is not universal and the color mapping might not be meaningful for some users.

For WheelSense:

- *Integrated control:* The control in the WheelSense systems (V2, V3 and V4) is integrated in the steering wheel, in order to be as close as possible to the driver's primary activity. In the WheelSense V1 approach, the control was not integrated in the object, but the affordances of the steering wheel helped the execution of the gestures.

- *Integrated representation:* For all the WheelSense systems, no feedback was provided inside the steering wheel, neither visual nor haptic feedback, in order to avoid distraction for the user. The WheelSense systems relied on environmental audio feedback from the media player and additional visual feedback was provided in WheelSense V4 through a Head-Up Display.

- *Direct control:* All gestures in the WheelSense systems had a direct effect within the media player, although the WheelSense V4 exploited a hierarchical menu to explore the different functions (such as playing music, or calling a contact). A more direct control would have required defining a different gesture for each function. Moreover, the volume control in WheelSense V4 can be performed through incremental steps: a more fine-grained and continuous control of the volume would have been preferable.

- *Direct representation:* For the media player, the information is represented by the different songs and, although intangible, is represented directly by system feedback. However, the hierarchical menu in the WheelSense V4 system introduces a level of abstraction to access the underlying information.

- *Meaningful control:* For the first three WheelSense systems, gestures were designed by experts, trying to exploit embodied metaphors to facilitate the associations between the gesture syntax and the gesture semantics. The WheelSense V4 relied on the intuitiveness of gestures elicited by users, which often exploited a spatial reference for the definition of gestures.

- *Meaningful representation:* As discussed before, the representation of the media player application was intangible, but still meaningful because it exploited inherent feedback from the media player. The representation of information in the WheelSense V4 system was also tied to the hierarchical menu, which was clear for the users but was still limited by the intangible text representation in the Head-Up Display.

The analysis of the ADA lamp tangible qualities shows that this system is compliant to all quality criteria but the direct control. However, in this system, the absence of direct control was designed on purpose to obtain an unpredictable, life-like behavior. Therefore, the lack of one or more tangible qualities is not always a synonym of a bad TGI system.

The Analysis of the WheelSense system showed that the representation of information was intangible, since the user cannot directly manipulate other objects than the steering wheel without safety risks. For this reason, while the control was integrated in the steering wheel, the system provided only intangible audio or visual feedback. In this context, it is worth noting that the WheelSense V2 system scored a higher result in terms of usability then the WheelSense V4 system 84 versus 66 points). Although these results were obtained in the context of two different experiences, where the WheelSense V4 interface provided several additional functions in respect to the WheelSense V2 interface, it could be argued that the hierarchical menu of the WheelSense V4 interface introduced an additional layer of abstraction that made the users rate better the more direct control allowed by the WheelSense V2 system.

From the analysis of the previous systems, it is evident the difficulty to evaluate the interaction qualities of a system with a binary (present/not present) rating. Sometimes a quality could be present only partially and the evaluation of that quality could depend on the social or personal background of the users. Moreover, the absence of a quality could be designed on purpose to adapt to the context of the application or to obtain particular characteristic of the interface (e.g., the unpredictability of the ADA lamp behavior, or the "eyes on the road, hands on the steering wheel" paradigm for WheelSense). In conclusion, the analysis of the tangible qualities of a TGI system could be useful to evaluate the properties of the system and to individuate possible routes for improvement, but it does not offer an absolute measure of its overall quality.

8.4 Generative Power

There is a large variety of new tangible gestures that can be explored and the aim of the move, hold and touch taxonomy is offering a large palette of gestures that can be performed in relation to virtually any object of the physical world. Combining gesture components (move, hold and touch) with attributes (pressure, amount, time, *etc.*), the designer can generate a plethora of different interactions with objects. Design principles like polymorphism and reification [24] (cf. Section 5.2) can be applied in a given application to generate new gestures. With the polymorphism, a multipurpose object can be transformed in different tools according to how it is held or moved (cf. [192, 217] and the smart pen in Section 7.6). This will reduce in the application the number of objects to interact with. The polymorphism principle allows also to use the same gesture to interact with different objects, having different effects in the application according to the object to which it is applied, as seen for the smart watch illustrated in Section 7.5 (cf. also the functional gesture approach of Carrino et al. [42]). New interactions can be generated also through the reification of digital concepts: indeed, it is possible to add existing objects in the real world as interactive tokens or to make the user define personalized gestures (cf. the Smart Watch application in Section 7.5). Finally, TGIF provides practical guidelines to build TGI systems, allowing applying the designed interactions to real world scenarios (Chapter 6).

The application examples provided in Chapter 7 explain how the guidelines proposed in TGIF have been used to generate seven TGI systems and the respective tangible gestures. Table 2 in Section 7.1 resumed how I explored the design space of TGIF through the application examples for three different application domains. The design space of TGIF explored through the systems presented in Chapter 7 is further depicted in Figure A10 in Appendix A. Obviously, the examples are not comprehensive, since the design space of TGIF is very broad. For example, none of the systems provided as examples used the environmental approach to recognize tangible gestures. Similarly, only a few of the common practices for application domains and object affordances has been applied. Nevertheless, it was possible to explore many different types of gesture syntax and semantics, different design approaches (expert-driven and user-driven) and three different technological approaches.

8.5 Conclusion

In this chapter, I presented a theoretical evaluation of the Tangible Gesture Interaction Framework, based on the Beaudouin-Lafon's definition [23] of a good interaction model. Practical evidence of the TGIF qualities has been shown within each criterion: for the descriptive power, I showed a classification of existing systems according to the proposed framework, as well as the results of an evaluation conducted with 4 HCI experts for the tangible gesture syntax; for the evaluative power, I pointed out to the different sections of the framework that provide guidance on the evaluation of TGI and of TGI systems and I proposed an analysis of two of the different applications according to the interaction quality criteria proposed in Section 5.3; finally, for the generative power, I pointed out to the guidance provided in the framework and I summarized the different systems developed in this thesis as practical applications of the framework.

As discussed in Section 8.1, validating frameworks, and in particular abstracting frameworks, is not an easy task. While a theoretical evaluation of TGIF based on criteria found in literature has been proposed, a practical evaluation of the framework with engineers and designers was not possible in the context of my thesis, for time and logistic constraints. Obviously, this practical evaluation would be very useful to spot further potential limitations of the framework and to improve it further; therefore, I hope that it will be possible to conduct such kind of evaluation as future work.

Part IV – Discussion and Conclusions

9 Discussion, Perspectives and Conclusions

"A torto si lamentan li omini della fuga del tempo, incolpando quello di troppa velocità, non s'accorgendo quello essere di bastevole transito; ma bona memoria, di che la natura ci ha dotati, ci fa che ogni cosa lungamente passata ci pare esser presente."
Leonardo da Vinci, "Aforismi, novelle e profezie".

In this chapter, I discuss the main achievements, as well as the limitations of the investigations conducted during this thesis and the perspectives for undergoing and future work.

9.1 Discussion of Thesis Contributions

The objective of this thesis, in the context of the project "Living in Smart Environments: Natural and Economic gesture-based HCI" supported by the Hasler Foundation, was the investigation of a novel interaction means that could enrich the experience of the user in smart environments. Tangible Gesture Interaction, a paradigm recently introduced by Hoven and Mazalek as the intersection between the more popular paradigms of tangible interaction and gesture interaction, has been chosen as the field of investigation of this thesis. At the beginning of this thesis, Tangible Gesture Interaction was a new research domain within the broader field of Human-Computer Interaction, blending the two popular domains of Tangible Interaction and Gesture Interaction. When I started my thesis in September 2010, Hoven and Mazalek had not published yet the first definition of Tangible Gesture Interaction [213], although that article was the starting point of my thesis. Besides Hoven and Mazalek's paper, any other explicit reference to this new domain was missing and there was a lack of specific guidance on this topic, especially on *how to support the creation of new systems that exploit tangible gestures as interaction means with computer*. Therefore, this was the starting point of this thesis and the main research question. Following these insights, Section 1.4 presented the main challenges of this thesis and the related research questions:

1. *How to formalize the syntax and the semantics of tangible gestures?*
2. *How to support the design of tangible gesture interaction?*
3. *How to support the implementation of tangible gesture interactive systems?*
4. *How to evaluate a tangible gesture interactive system?*

These questions have been answered through the definition of the Tangible Gesture Interaction Framework (TGIF), presented in Chapters 4, 5 and 6 of this thesis. These contributions are discussed in detail in Section 9.1.1.

Moreover, seven different tangible gesture interactive systems have been designed and developed as applications of the proposed framework. Each system brought particular contributions to the respective application domain and allowed to discover peculiarities of TGI systems. These contributions are discussed in detail in Section 9.1.2.

9.1.1 Contributions of the Tangible Gesture Interactive Framework and Discussion

The contributions of the Tangible Gesture Interaction Framework will be discussed according to the four questions presented in Section 1.4 and reported at the beginning of this Section.

- *How to formalize the syntax and the semantics of tangible gestures?*

While Hoven and Mazalek provided extensive examples for the characterization of tangible gesture interaction [213], no formalization was provided for describing which forms tangible gesture can assume and which semantic constructs can be used to associate meaning to tangible gestures. In order to overcome this lack and to answer this research question, first, tangible gesture interaction has been modelled as a language, defining a communication model based on tangible gestures as signs to communicate with computer (Section 4.2). Second, this thesis proposed a taxonomy for describing the syntax of tangible gestures, which are considered as gesture-object pairs and composed by three optional components (move, hold and touch) and additional attributes that are performed in relation to one object (Section 4.3). Third, this thesis proposed a taxonomy for the classification of semantic constructs that can be used to associate meanings to gesture-object pairs. In particular, this taxonomy classifies existing constructs found in literature according to the level of coherence between the real and digital world in a two-dimensional gesture-object continuum (Section 4.4).

The aforementioned contributions have been validated classifying existing tangible gesture interactive systems (Section 8.2.1) and the tangible gesture interactive systems developed during this thesis (Section 8.4). Finally, an additional evaluation has been conducted with 4 HCI experts for a preliminary version of the TGIF syntax (Section 8.2.1).

In conclusion, the abstracting part of TGIF (Chapter 4) aims at providing a large palette of tangible gestures to be adopted in interactive systems as well as a palette of semantic constructs that can be used to associate meanings to tangible gestures, helping designers reflecting about all the possible gesture interactions with objects.

- *How to support the design of tangible gesture interaction?*

In order to answer this research question, first, the typical process for designing a tangible gesture interactive system has been discussed (Section 5.1), highlighting the importance of continuously iterating and assessing the result of the design (Section 5.7). Moreover, the framework suggested two typical design methods (expert-driven or user-driven) to design gesture with objects (Section 5.2). To support expert-driven design methods, common practices found in literature for designing tangible gesture interaction for different application domains (Section 5.4) and the typical object affordances that can be exploited to support particular tangible gestures have been summarized (Section 5.5). Finally, the typical feedback that can be designed for tangible gesture interactive systems has been discussed, evidencing the two main types, i.e., continuous feedback during gesture execution for manipulative gestures and discrete feedback at the end of the gesture for command-like gestures (Section 5.6).

These contributions have been validated through the design of 7 tangible gesture interactive systems for three different application domains, following different design methods and common practices, or exploiting different object affordances (Chapter 7, cf. also Section 8.4).

In conclusion, while the abstracting part of TGIF offered a broad overview on tangible gesture forms, there were no particular hints about how these gestures can be chosen and designed for interactive systems. To this purpose, the designing part (Chapter 5) proposed an overview of the process for conceiving a tangible gesture interactive system and of the different methods for designing tangible gestures.

How to support the implementation of tangible gesture interactive systems?

In order to answer this research question, first, four different approaches for the recognition of tangible gestures have been presented, i.e., embedded and embodied, wearable, environmental and hybrid, discussing the typical advantages that each approach can bring in a tangible gesture interactive system according to eight design parameters and the typical sensors that can be used in each approach to recognize tangible gestures (Section 6.1). Second, this thesis summarized typical hardware platforms that can be used to build embedded and embodied or wearable tangible gesture interactive systems, illustrating how to choose the most appropriate platform according to 7 different parameters (Section 6.2). Finally, typical challenges and software solutions for the recognition of tangible gestures have been presented, with a particular focus in gesture segmentation, gesture classification, and fusion techniques for multi-classifier systems (Section 6.3).

These contributions have been also validated through the development of 7 tangible gesture interactive systems using three different technological approaches, i.e., embedded and embodied, wearable and hybrid (Chapter 7, cf. also Section 8.4).

In conclusion, to support the actual development of tangible gesture interactive systems, the building part of TGIF (Chapter 6) suggested four different approaches for recognizing tangible gestures and provided generic guidance for the hardware and software that can be used to build TGI systems.

- *How to evaluate a tangible gesture interactive system?*

In order to answer to this research question, the evaluation criteria for the quality assessment of tangible gesture interaction (Section 5.3) and typical methods for evaluating the objective performance of tangible gesture interactive system (e.g., recognition accuracy) and the subjective experience of the user (e.g., usability of the system) (Section 5.7) have been presented. This contribution has been also validated conducting different evaluations for the systems described in Chapter 7. Moreover, an evaluation of the interaction qualities of two TGI systems has been presented in Section 8.3.1. In conclusion, although TGIF does not offer a dedicated Section for the evaluation of tangible gesture interactive systems, specific guidelines are still provided in Chapter 5.

From the analysis of the contributions to the specific research sub-questions, it is worth discussing the contribution to the main research question: *how to support the creation of new systems that exploit tangible gestures as interaction means with computer?*

As a consequence of the previous analysis of the thesis contributions, the proposition of this thesis is to support designers throughout the whole process of design and development of tangible gesture interactive systems. This has been concretized with the definition of the Tangible Gesture Interaction Framework (TGIF), which helps designers, first, to reflect about the nature of tangible gesture interaction and the forms that tangible gestures can assume. Then, more in practice, it helps designers during the different interaction design and development phases of the system, through a series of guidelines and best practices to be used as a reference to explore the large design space of tangible gesture interaction. The usefulness of the framework for other researchers need still to be validated, however, a formal validation of the framework has been already presented in Section 8.

After answering the main research question, it is still worth discussing about the importance of this work in the context of Human-Computer Interaction. Indeed, although the specific domain of tangible gesture interaction was recently introduced, much knowledge provided in this framework is based on previous work from tangible interaction and gesture interaction. Since gestures have always been used in tangible interaction and gestures have often been designed for interacting with smart objects or smart environments, one could ask which is the novelty of tangible gesture interaction. After the different experiences conducted during this thesis, in my opinion, the novelty of tangible gesture interaction is that it cares about objects as much as gestures, considering them as a whole, i.e., as tangible gestures or gesture-object pairs, while the two existing research fields focus either on objects and how they can be manipulated or in the expressivity of gestures and how they can be recognized. Meanwhile, in these last years, tangible interaction is broadening its perspectives to find richer experiences for the user and, at the same time, there is an increasing interest in applying gesture interaction to the interaction with the physical world. TGI foundations are grounded at the intersections of these concurrent needs.

While in Section 2.4 I have mentioned some previous frameworks that have a similar holistic perspective on gestures with object, they were limited to particular types of gestures or application domains. This thesis tried to frame the whole design space of gestures with objects with a broader perspective, which is probably ambitious to cover within a single thesis. For this reason, some elements of the framework have been treated only superficially, but this broader perspective could be valuable for a researcher or practitioner that approaches this field for the first time (as I was at the beginning of this thesis and as I felt missing for my research). As discussed in Section 8.1, assessing how much valuable TGIF is for designers or developers of tangible gesture interactive systems was particularly difficult in the context of this thesis. An effective adoption of the framework can be assessed only after several years.

Considering the starting point of this thesis, i.e., the purpose of bringing back human-computer interaction to the real world, following the manifest of Wellner et al. [223], tangible gesture interaction perfectly fits this purpose. This is particularly evident from the practical applications that have been developed during this thesis. The different systems for gesturing on the steering wheel tried to embrace the paradigm “eyes on the road, hands on the steering wheel”, exploiting the affordances of the steering wheel for reducing the driver’s distraction, helping him focusing on the primary task. The ADA lamp was designed with the purpose to have more natural emotional exchanges at distance, exploiting the anthropomorphic affordances of the lamp and bringing the interaction beyond the screen. The Hugginess system is probably the maximum expression of this intent, encouraging people to hug in physical encounters instead of spending time looking at their phones. The smart watch was also a way to exploit the physical world and the objects that surround the user at home, exploiting them as tangible cues to memory and allowing a ubiquitous interaction thanks to a wearable system that accompanies the user on the move.

Within the context of this thesis, i.e., the project “Living in smart environments”, with the particular purpose of looking for a new natural and economic language to interact in smart environments, the applications presented in Chapter 7 explored the possibility to use tangible gestures as a communication means with computer in different smart environments, in particular at home and in a vehicle (and marginally at work with the smart pen system). This thesis has shown that tangible gestures can scale as a language and the semiotic analysis conducted in Chapter 4 reflects this assumption. Naturalness of gestures have been largely discussed in HCI [147] and Section 1.1 presented the different definitions that exist of natural interfaces. In particular, tangible gesture interaction can be considered as a natural interface since it exploits the innate human perceptive-motor skills and cognitive skills. Benefiting from the human ability to manipulate objects and to build knowledge from the exploration of the physical world [216], it is possible to design intuitive interactions either with physical artifacts built on purpose for the desired system or with existing objects of our surrounding world.

Conversely, the economical aspect of tangible gesture interaction could not be granted in all interaction scenarios. As shown in Section 7.5.5 for the smart watch experience, users preferred interacting with physical artifacts without holding them in the hand, which can be tiring in some cases. Indeed, the well-known gorilla arm-problem of gesture interaction can be accentuated in tangible gesture interaction if the users have to perform gestures with a heavy object or with an object that is difficult to hold. Designing economical interactions in TGI is obviously still possible, since the designer can choose among a large variety of different tangible gestures, adapting them to the contexts of the different applications. Finally, the economical aspect of the interaction can be discussed from a cognitive point of view. The possibility to integrate seamlessly tangible gestures in the users' routines depends mostly on the ability of the designer to exploit the users' peripheral attention, the physical and cognitive affordances of the objects and the intuitiveness of gestures (cf. WheelSense in Section 7.2). However tangible gestures can be designed to make the users reflect, focusing their attention on the gesture performance, which can assume a particular relevance also for other people and not only for the machine (cf. Hugginess in Section 7.4).

To conclude this Section, I point out to the publications that presented the theoretical contributions of the thesis: [12, 13].

9.1.2 Contributions Obtained in the Different Application Domains

In addition to the theoretical contribution brought by the Tangible Gesture Interaction Framework, specific contributions have been achieved for the different interactive systems developed during this thesis. In particular:

- WheelSense V1 (Wearable): This is the first system that recognizes gestures (4 classes) performed while holding the steering wheel with a wearable approach. The classifier scored an overall recognition accuracy of 94.55% on pre-segmented gestures. (Section 7.2.2)

- WheelSense V2 (Embedded and embodied): This is the first system that recognizes gestures (4 classes) performed while holding the steering wheel through pressure sensors distributed on its surface. The classifier scored an overall recognition accuracy of 82% on gestures explicitly segmented by the user. The overall usability of the system was 84 points. (Section 7.2.3)

- WheelSense V3 (Hybrid): This study introduced a novel paradigm, called opportunistic synergy, to take advantage of heterogeneous systems for tangible gesture interaction and combine them in terms of system performance (recognition accuracy), user experience and resources. The study compared 8 different classifier fusion techniques (some from literature, the others designed on purpose for this study), two segmentation strategies (implicit and explicit) and 3 different segmentation fusion techniques for the recognition of five different gestures performed on the steering wheel and recognized through a hybrid system obtained by the combination of the WheelSense V1 and V2 systems. The best results (F1= 85%, Accuracy=86%) have been obtained with the MCC* and MCC+* classifier fusion techniques, using a manual segmentation strategy with an ADAPTIVE segmentation fusion technique. (Section 7.2.4)

- WheelSense V4 (User-Driven): This is the first system that adopted a gesture elicitation approach to individuate tangible gestures performed on the whole surface of the steering wheel as an interaction means with the In-Vehicle Infotainment System. The gesture elicitation study (Section 7.2.5) allowed to derive a taxonomy of gestures that users are more likely to perform on the steering wheel, either releasing the hands from it or while firmly grasping it. The study allowed also to demonstrate that participants that were not instructed to elicit gestures for predefined commands produced a larger variety of gesture types and performed gestures with more different body parts. Moreover, the study allowed individuating the steering wheel zones that are more suitable to perform gestures on it. The WheelSense V4 system, developed choosing the 6 most popular gestures obtained from the gesture elicitation, has been compared with two other systems that exploited, respectively, a vocal interface and a touch interface in the central dashboard, in order to estimate the differences in terms driving performance, secondary task performance and user experience. The only significant differences found were limited to the secondary task performances (the touch interface having a task completion time inferior to the WheelSense interface, and the vocal interface requiring a fewer number of interactions than the other two interfaces). (Section 7.2.6)

- ADA Lamp: This system proposed the design of the first anthropomorphic lamp for emotional communication at distance. The system proposed a novel design to integrate anthropomorphic affordances and a life-like behavior in the lamp and a low-tech solution for recognizing 5 gestures performed on the lamp surface and typically used for human-to-human interpersonal communication. A system that simulated a life-like companion reacting to users' gestures has been developed. (Section 7.3)

- Hugginess: This is the first system to augment hugs between people during social encounters and to encourage them through a dedicated smartphone app. This is also the first wearable system that implemented a digital information exchange between the fabrics of the t-shirts of two hugging people. The project allowed also developing the first hug counter for free-hug events. (Section 7.4)

- Smart Watch: This project proposed the first design of a smart watch to recognize gestures performed with objects in the hand. In particular, it allowed conducting a feasibility study to embed a RFID reader in the smart watch and recognize tagged objects held in the hand, and proposed the first solution to recognize tendon movements generated by different grasps, finger gestures and wrist flexion with pressure sensors embedded in the watch strap. Through a Wizard of Oz experience, it allowed also to investigate user preferences about gesturing with objects for controlling a media player in a smart environment. (Section 7.5)

To conclude this Section, I point out to the publications that presented the practical contributions of the thesis: [4, 5, 7-11, 41, 126, 142].

9.2 Limitations

The goal of this thesis was quite ambitious, since a framework for a new interaction paradigm that blends concepts from two existing disciplines opens the way to several research questions. Since no previous framework for tangible gesture interaction existed, TGIF aims at framing past and future work in this field only at a high level. The main objective of TGIF was demonstrating the richness of tangible gesture interaction, without offering a formal grammar for describing all the possible tangible gestures. Such grammar could facilitate the definition of more formal descriptions of tangible gestures and this description can be used for their recognition of tangible gestures in the interactive systems, as shown for surface gestures (cf. Section 2.3.3.3). However, because of the variety of tangible gestures that can be imagined, a generic grammar for all tangible gestures will be very complex, with the result of confusing novice designers who approach this field for the first time. Moreover, since TGIF can be applied to several application domains, it was not possible to provide detailed guidelines for each of these domains. For the same reason, only three specific application domains have been explored with practical applications.

The provided syntax for tangible gestures does not provide sufficient expressive power to describe every tangible gesture: as discussed in Section 8.2.2, gestures that focus on object movements or transformations are difficult to describe with the provided syntax without a shift of the point of view from which tangible gestures are analyzed (from a user-centered perspective to an object-centered perspective). Moreover, the TGIF syntax provides no support for the analysis of gesture performed with multiple objects.

TGIF does not address the very specific need of each TGI designer, but can help them pointing out to more specific guidelines, whenever they exist for a specific type of tangible gesture, application domain or technological approach for tangible gesture recognition. Nevertheless, the suggested common practices, which analyze a limited number of the numerous potential application domains of TGI, are not necessarily the best choice for all TGI systems: designers should always explore new gestures and new technologies in order to advance the specific field that they are investigating. Other aspects of tangible gesture interaction design have been discussed only superficially: feedback can assume various forms that I could not discuss in depth in this thesis.

Similarly, for the building section, the environmental approach for the recognition of tangible gestures has not been investigated in this thesis.

As already discussed in Section 8.1, the assessment of the framework has been done only on a theoretical level, since a practical assessment would have require additional time once the framework was completed as well as the possibility to work closely with designers and system integrators willing to use TGIF as guidance for their work. In addition, some of the applications developed during this thesis have not been assessed in depth. Besides the obvious time constraint, some systems could not be assessed with users because of their level of completion. For example, the development of a complete prototype of a smart watch would have taken a consistent effort, especially in terms of electronic integration, which was out of the scope of this thesis.

9.3 Perspectives

As discussed in Section 9.2, some aspects of the framework need additional investigations and assessment. Concerning the abstracting part of the framework, it would be interesting to define a formal grammar for describing tangible gestures and, for building part, to develop a toolkit that can automatically recognize the gestures defined through such grammar. As discussed before, probably it is not possible to define a complete grammar for all the variety of tangible gestures that can be designed. As a result, the associated toolkit would be probably limited to a subset of all the possible interactions. Still, such a toolkit would be helpful for the rapid development of basic tangible gesture interactive systems, which is useful for all the situations in which the designers do not have the time or the necessity to explore novel interactions. In the case that such interactions would be adopted in different application domains, as suggested by Prof Lorenzo Sommaruga, co-examiner of this thesis, it would be interesting to assess the interoperability of the designed interactions.

Recently, at the TEI'16 conference an article of Petrelli et al. [159] inspired further research for TGIF. The article investigated the aesthetic perception of the users facing objects with different shapes (spheres and cubes), textures (fabric, plastic), size (holdable with 1 hand / 2 hands) and behavior when picked up (visual, aural, haptic or no feedback), i.e., 32 different interactive objects. The purpose of the study was to assess how each object was perceived by the user according to several aesthetic criteria: Interesting, Comfortable, Playful, Surprising, Pleasant, Special and Relaxing. To evaluate the objects, participants had to hold them in the hand and the object delivered feedback when picked up. The study allowed determining that people generally prefer spheres to cube and fabric materials to plastic materials, as well as that objects with visual feedback (a blinking LED) were preferred to objects with no feedback when picked up. These results are particularly relevant also for TGIF, since they provide generic user preferences for the object design, independently from the application domain. Within this context, it would be interesting to investigate which gestures would be preferred by the users when performed in relation to objects with different characteristics. Indeed, while the Petrelli et al.'s study investigated only the hold gesture, assessing the users' aesthetic perception of different gestures such as touch, touch+move, move, hold+move and hold+pressure could provide valuable results for TGI. As for the previous study, it would be interesting also to discover if users' preferences vary according to the object shape, material, size and behavior. The results of such study would be useful to provide generic guidelines for tangible gesture design, independently from the application domain.

Additionally, as suggested by Prof. Patrizia Marti, co-examiner of this thesis, it would be interesting to study the coupling between input and output, i.e., between tangible gestures and system feedback. During the analysis of previous work, I could not notice any particular correlation between tangible gesture components and feedback types. Nevertheless, it would be worth exploring the expressivity of the different components and attributes, especially when directly coupled with different feedback types in a continuous manner (see Section 5.6). The framework of Wensveen et al. [224] could be source of inspiration for designing feedforward and feedback that are coupled directly to tangible gesture execution. However, further experiences and explorations are needed for tangible gestures, possibly following the same approach of Petrelli et al. [159].

A particular application domain for tangible gestures that is worth investigating is the Internet of Things (IoT). While this thesis already investigated the interaction with smart environments or with smart objects, the Internet of Things has particular characteristics that make the interaction with such objects particularly challenging. Indeed, since IoT objects are connected to the Internet and can operate autonomously, often it is difficult for the user understanding and controlling their behavior. Moreover, most current interfaces for IoT devices rely on smartphone or web apps, reducing the ability to grasp the behavior of such devices from the physical world. Nevertheless, IoT devices permeate our everyday environment, they integrate computational power and sensors and seem the perfect devices to start bringing users' interactions in the physical world, exploiting human's innate skills. In a context in which the number of such devices is increasing continuously, applying tangible gesture interaction to IoT objects seems a promising research field, with many potential applications in real case scenarios.

Finally, there is another interesting aspect that is worth discussing in relation to tangible gesture interaction and more in general to the design of interaction with real world objects: the possibility to make your own interactive objects. The maker movement is nowadays a reality and the large availability of 3D-printers to build custom objects and the large diffusion of prototyping toolkits such as Arduino (cf. Section 2.2.2 and 6.2) are bringing the possibility to build physical interactive systems to almost every person with minimal technical skills, beyond professional and researchers. Making your own interactive system brings a double benefit: first, the maker, which is often also the user, has complete control over the system and can adapt the system to his or her needs, or repair it whenever the system fails; second, the making experience allows to create an emotional attachment with the developed interactive system, especially if the system is embodied in a physical object, crafted by the maker. To support this latter consideration, during the development of the systems presented in Chapter 7, with a particular emphasis for the ADA lamp system, I could perceive this feeling of emotional attachment and the sentiment that the system behavior (especially its random malfunctioning) reflected a sort of soul of the object. In the case of the ADA lamp, the designed life-like behavior contributed to the feeling of being attached to a creature more than to an artifact.

Although the previous considerations could seem limited to the relatively small community of makers, they should make reflect to the impact that they could have on our increasingly consumerist society. Indeed, on the opposite side of the maker community, we find big industrial players that launch each year several new interactive devices, which are massively entering our homes. Such devices, having little or no personalization and built with short-lived disposable hardware and almost no repair possibilities (to allow selling new devices in the next years), they will hardly allow users to build long-lasting interactions and emotional attachment. As a result, we should reflect on the ecological and ethical impact that such impoverished devices could bring to our society.

In conclusion, whether tangible gesture interaction will be adopted by makers or by the industry to design new interactive objects, the hope is that it will be used to embed values and create stronger emotional attachments with objects, offering long-lasting interactions that better integrates with our natural real-world life, eventually enabling the possibility to readapt or recycle the device, instead of creating mere disposable gadgets that as only impact on our lives they will make grow the garbage stack after just few months of utilization.

Appendix A

This Appendix contains additional material that contributes to the illustration of the TGIF framework and of the application developed during this thesis. In particular, it presents several different schemas that shows the design space of TGIF and how the different system presented in Chapter 7 contributed to the exploration of this design space.

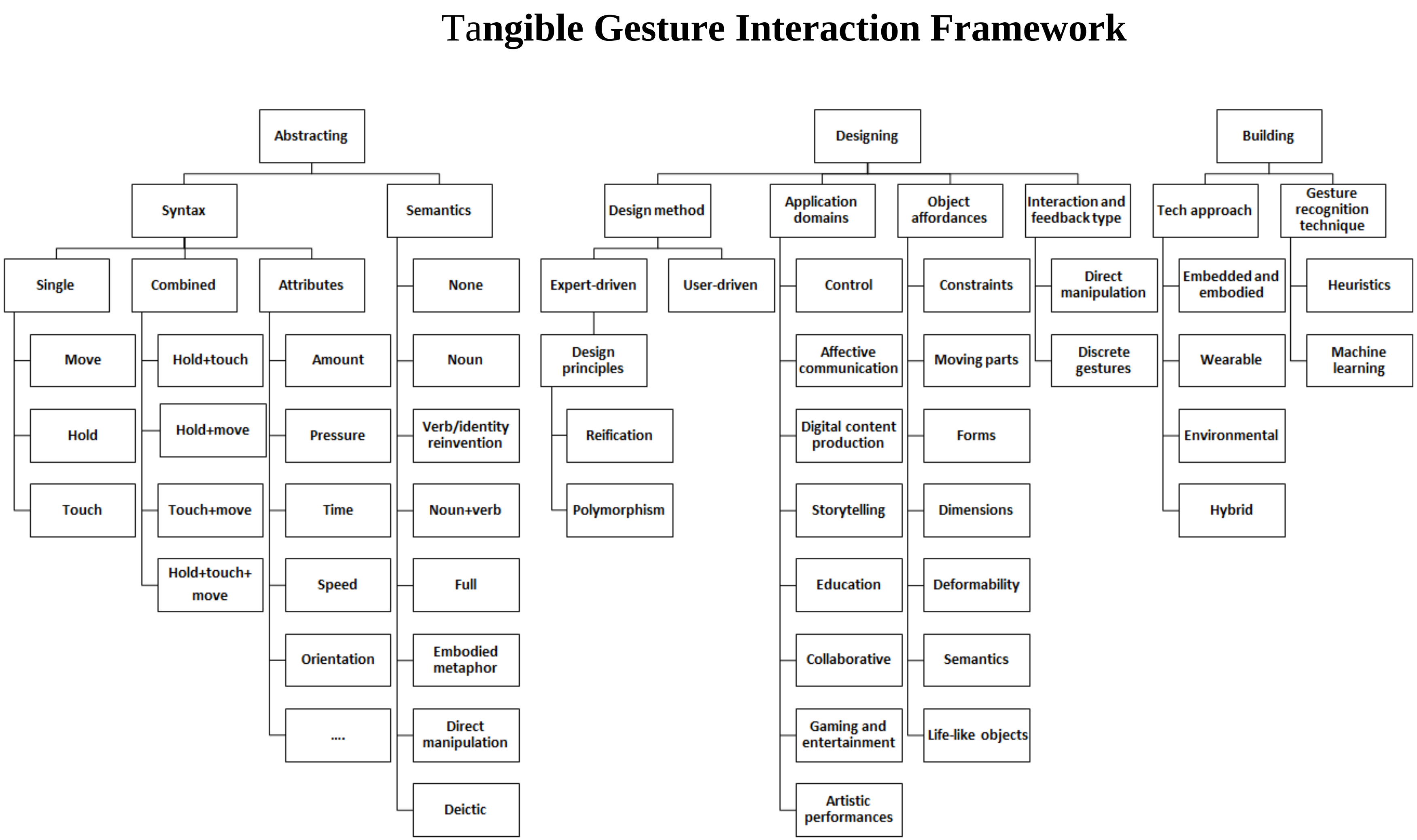

Figure A1. Design space of the Tangible Gesture Interacion Framework

WheelSense V1 – TGIF Exploration

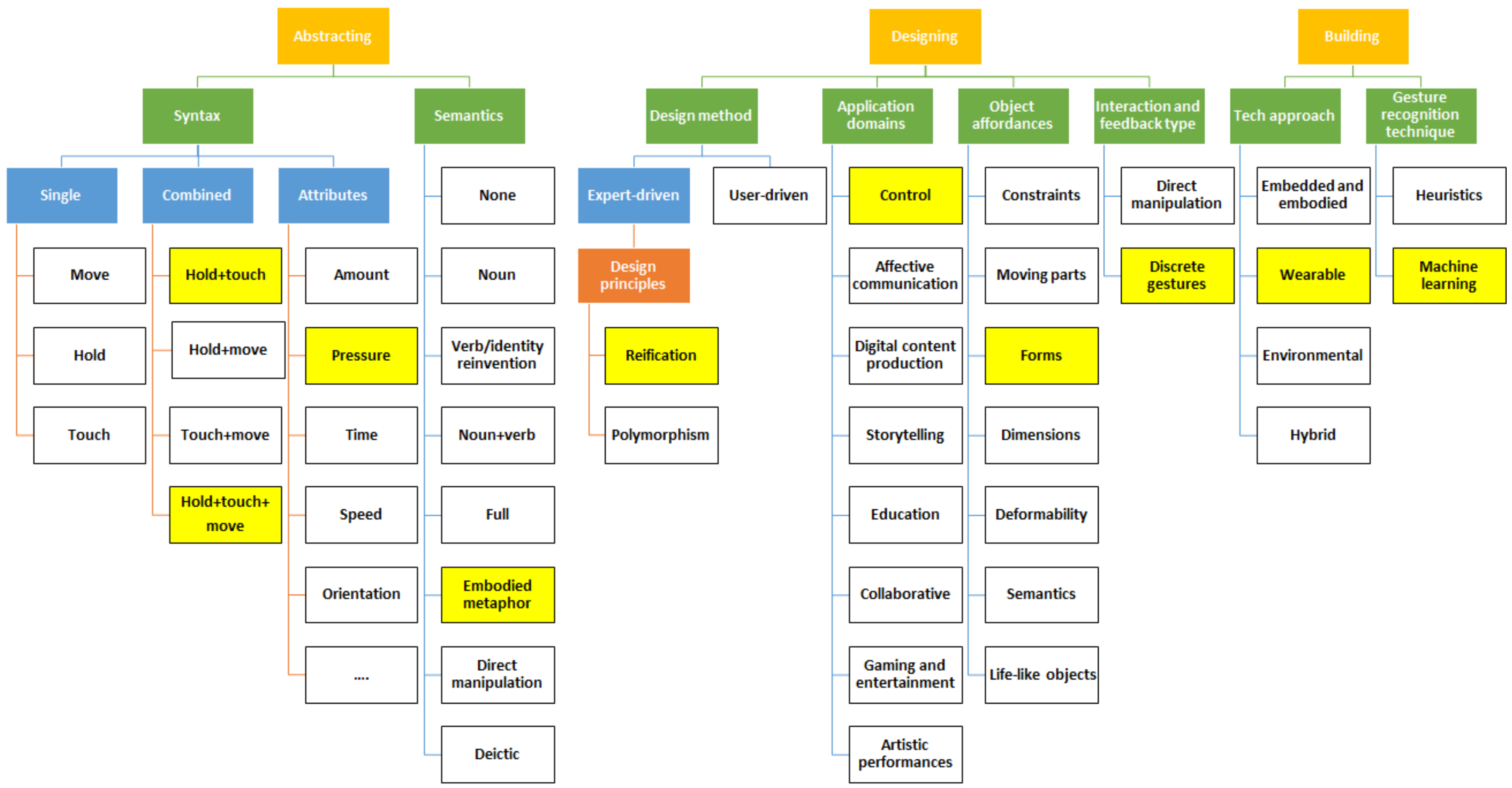

Figure A2. Elements of the TGIF design space explore in WheelSense V1

WheelSense V2 – TGIF Exploration

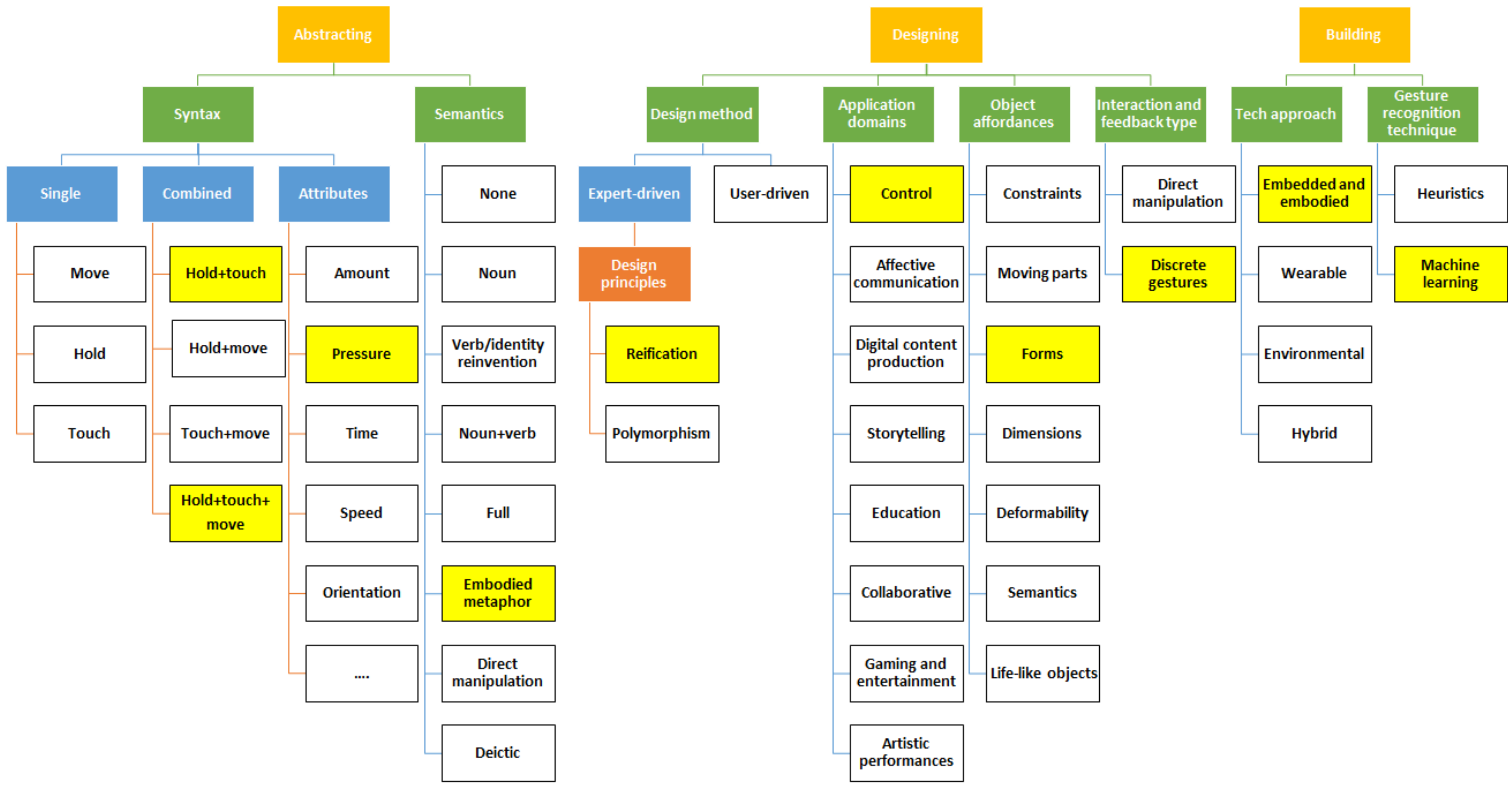

Figure A3. Elements of the TGIF design space explore in WheelSense V2

WheelSense V3 – TGIF Exploration

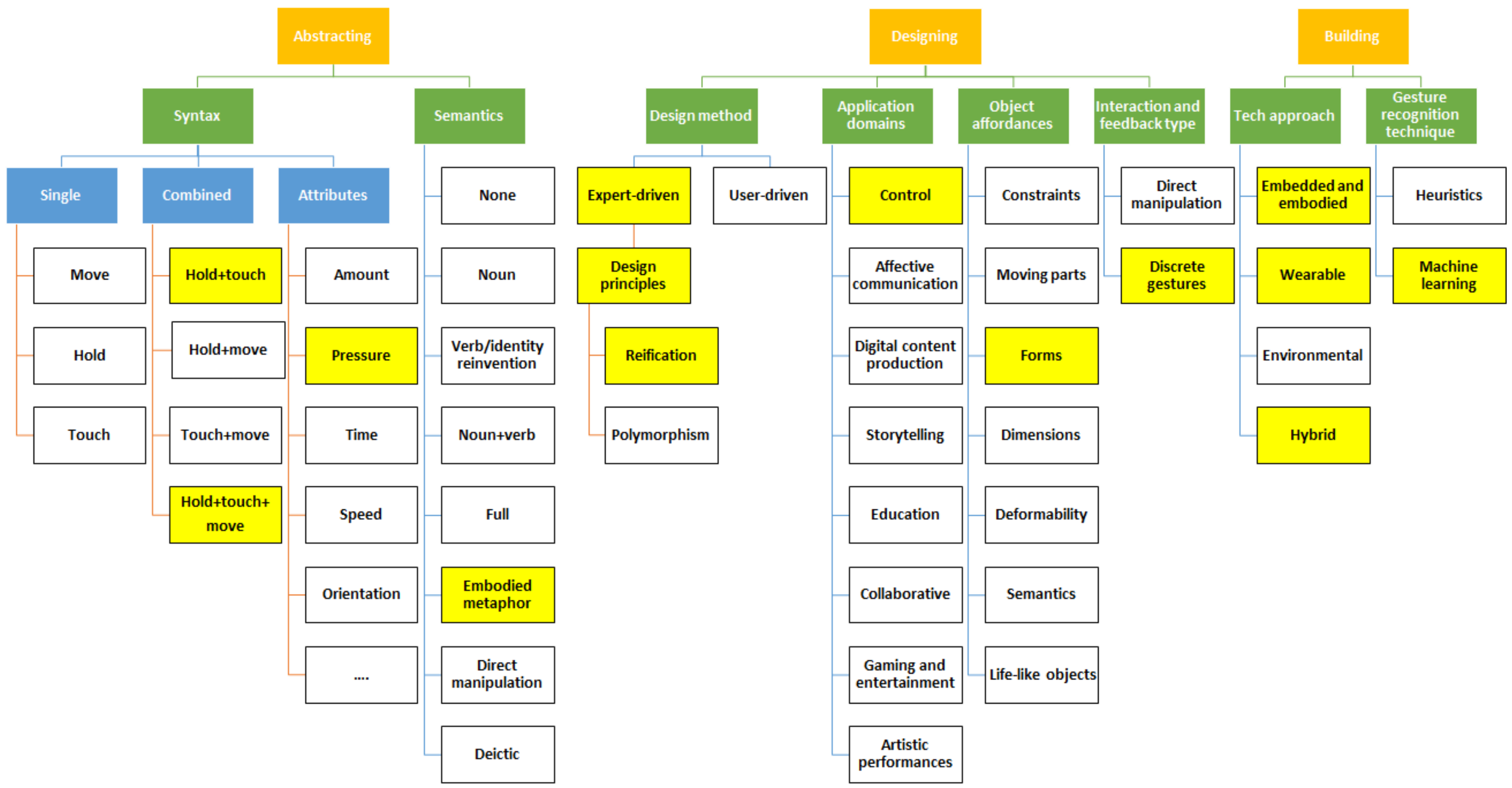

Figure A4. Elements of the TGIF design space explore in WheelSense V3

WheelSense V4 – TGIF Exploration

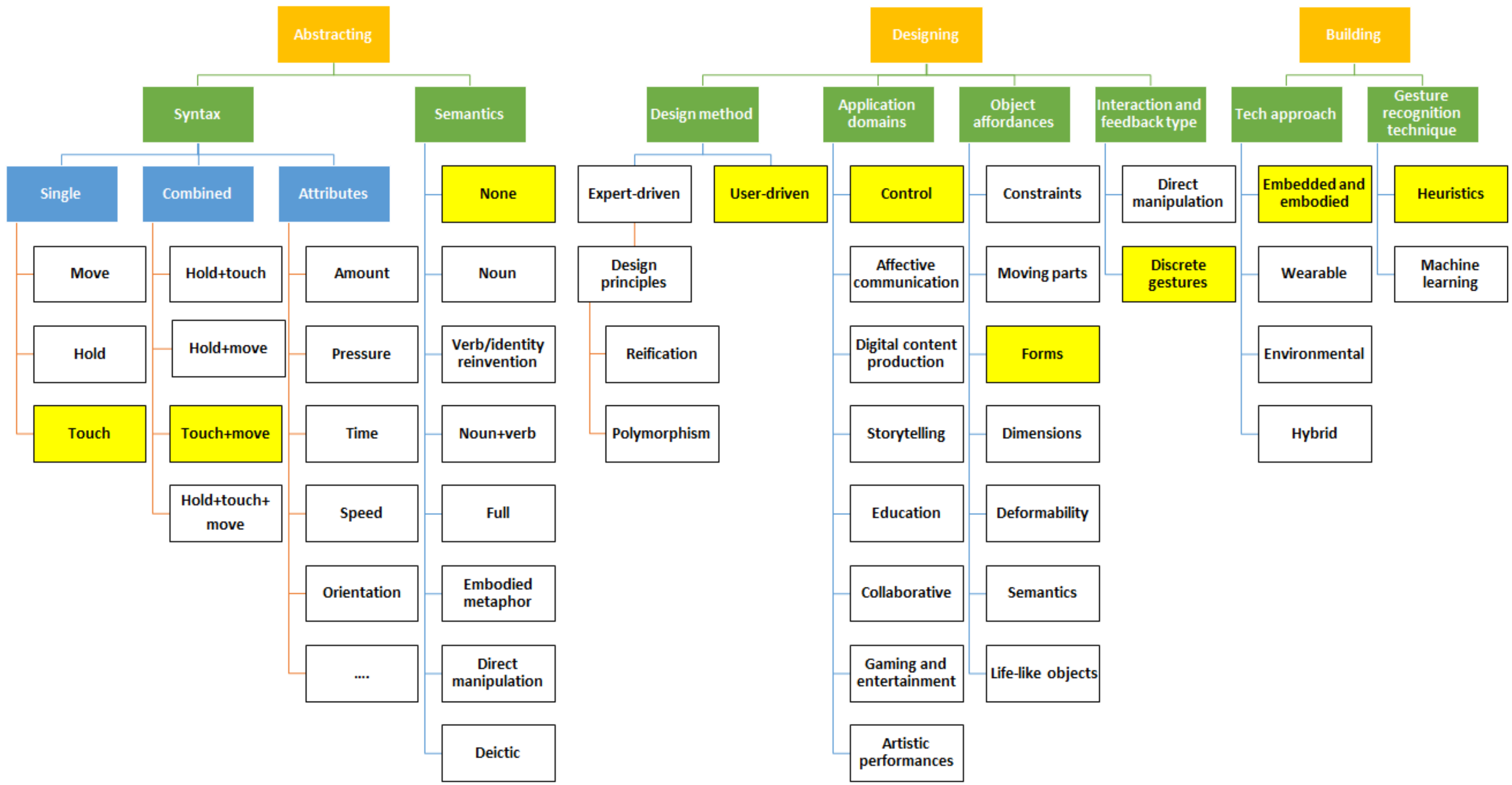

Figure A4. Elements of the TGIF design space explore in WheelSense V4

ADA Lamp – TGIF Exploration

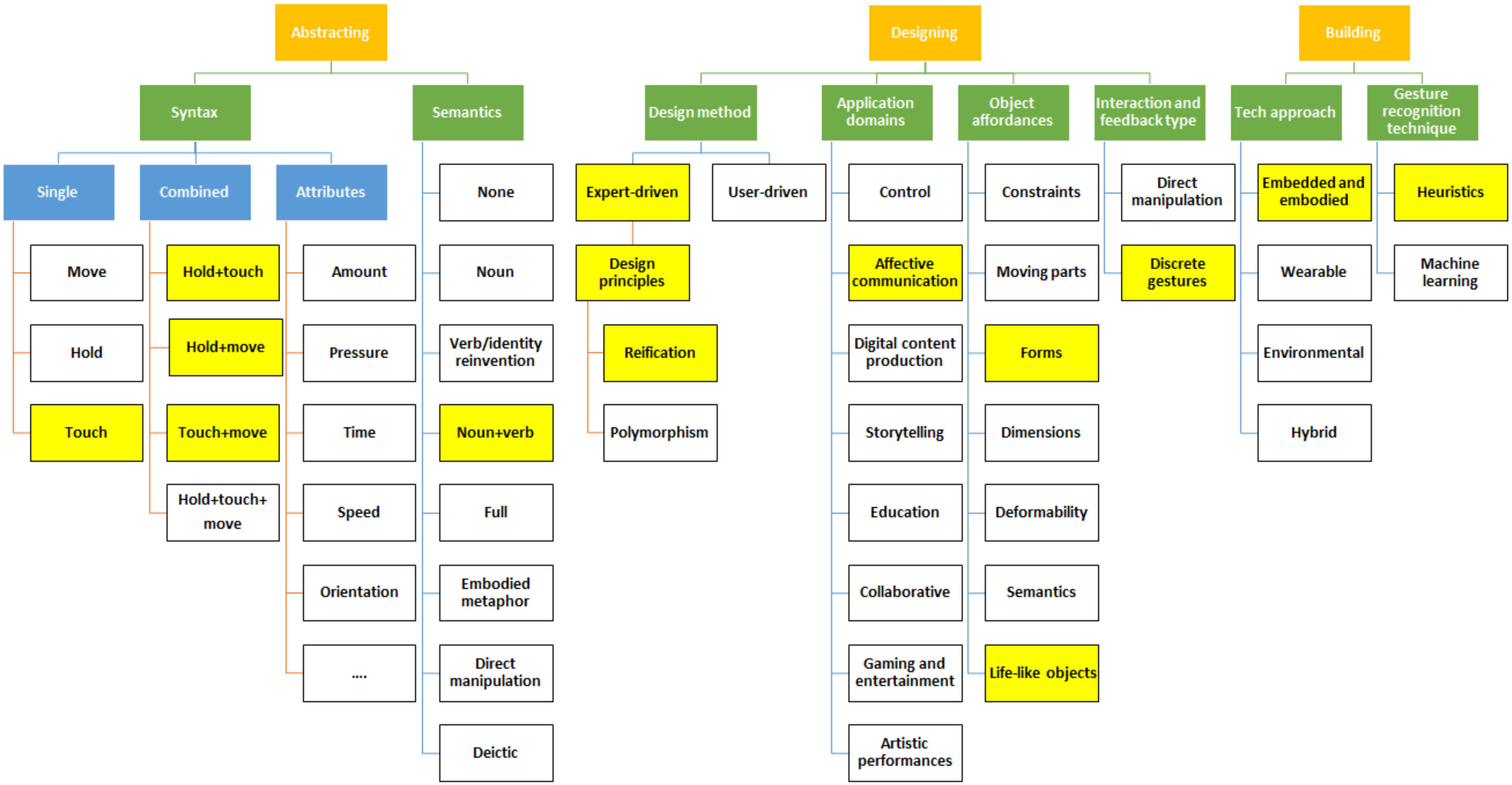

Figure A6. Elements of the TGIF design space explore in ADA Lamp

Hugginess – TGIF Exploration

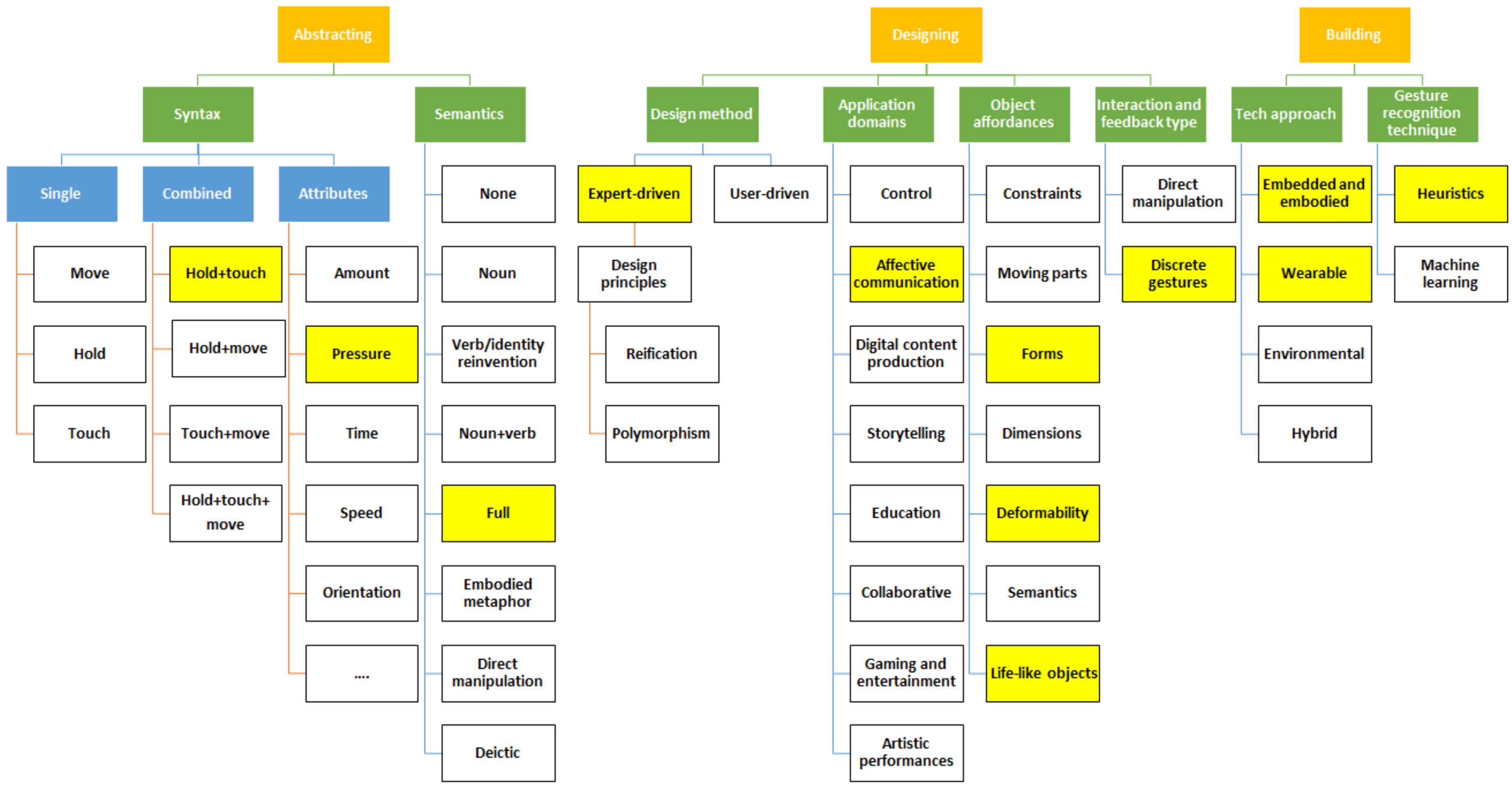

Figure A7. Elements of the TGIF design space explore in Hugginess

Smart Watch – TGIF Exploration

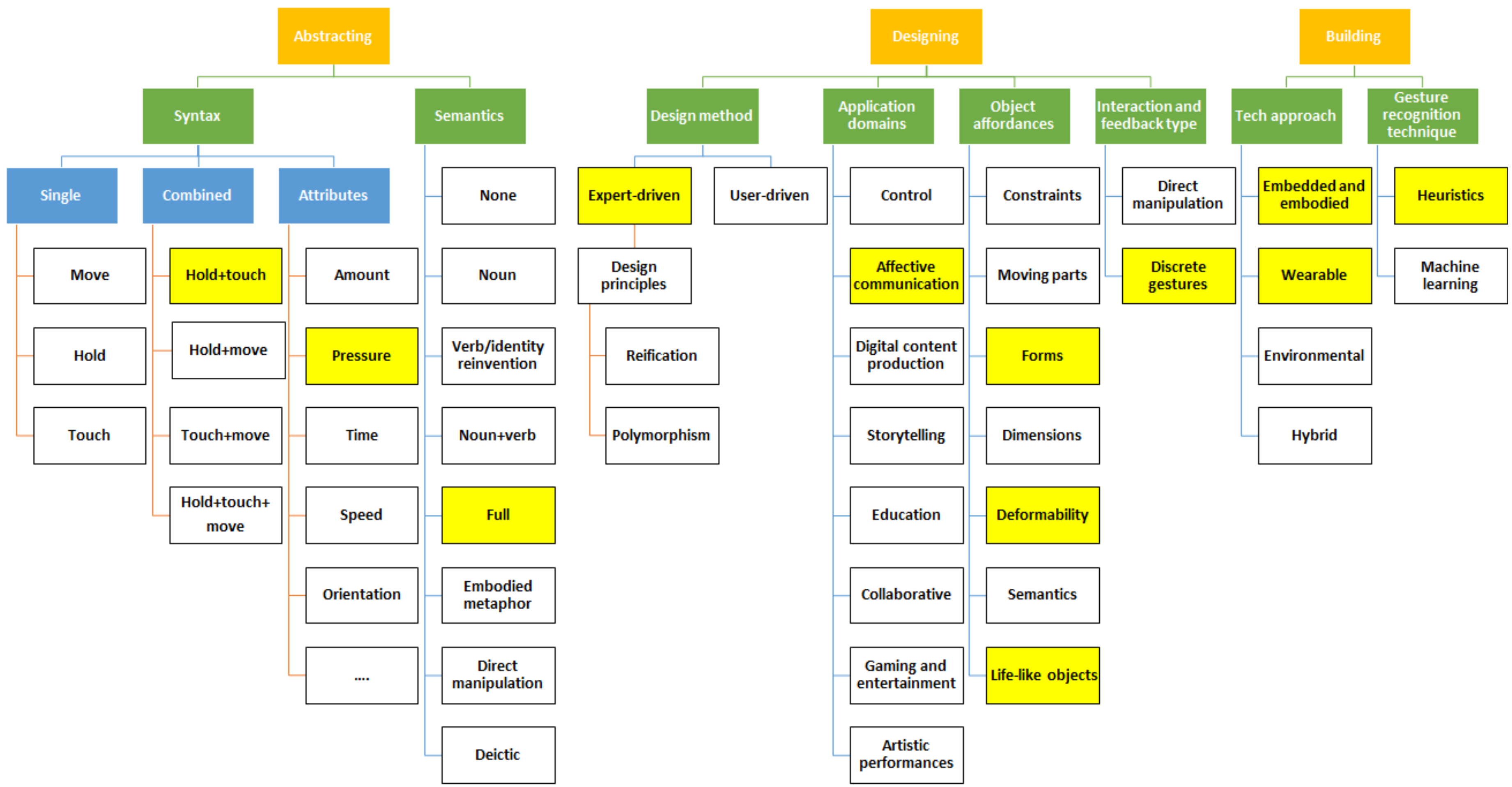

Figure A8. Elements of the TGIF design space explore in Smart Watch

Smart Pen – TGIF Exploration

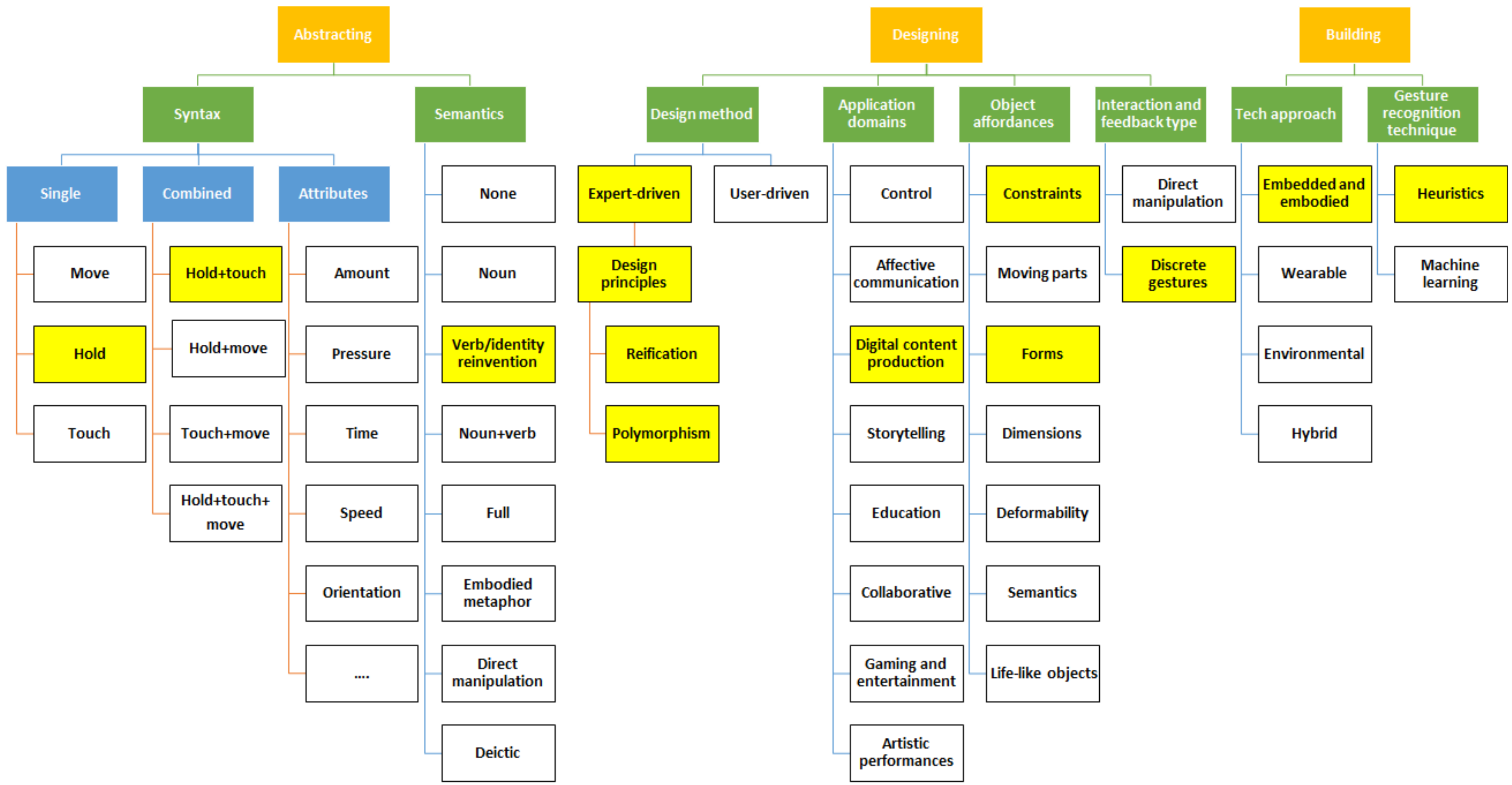

Figure A9. Elements of the TGIF design space explore in Smart Pen

TGIF Overall Exploration

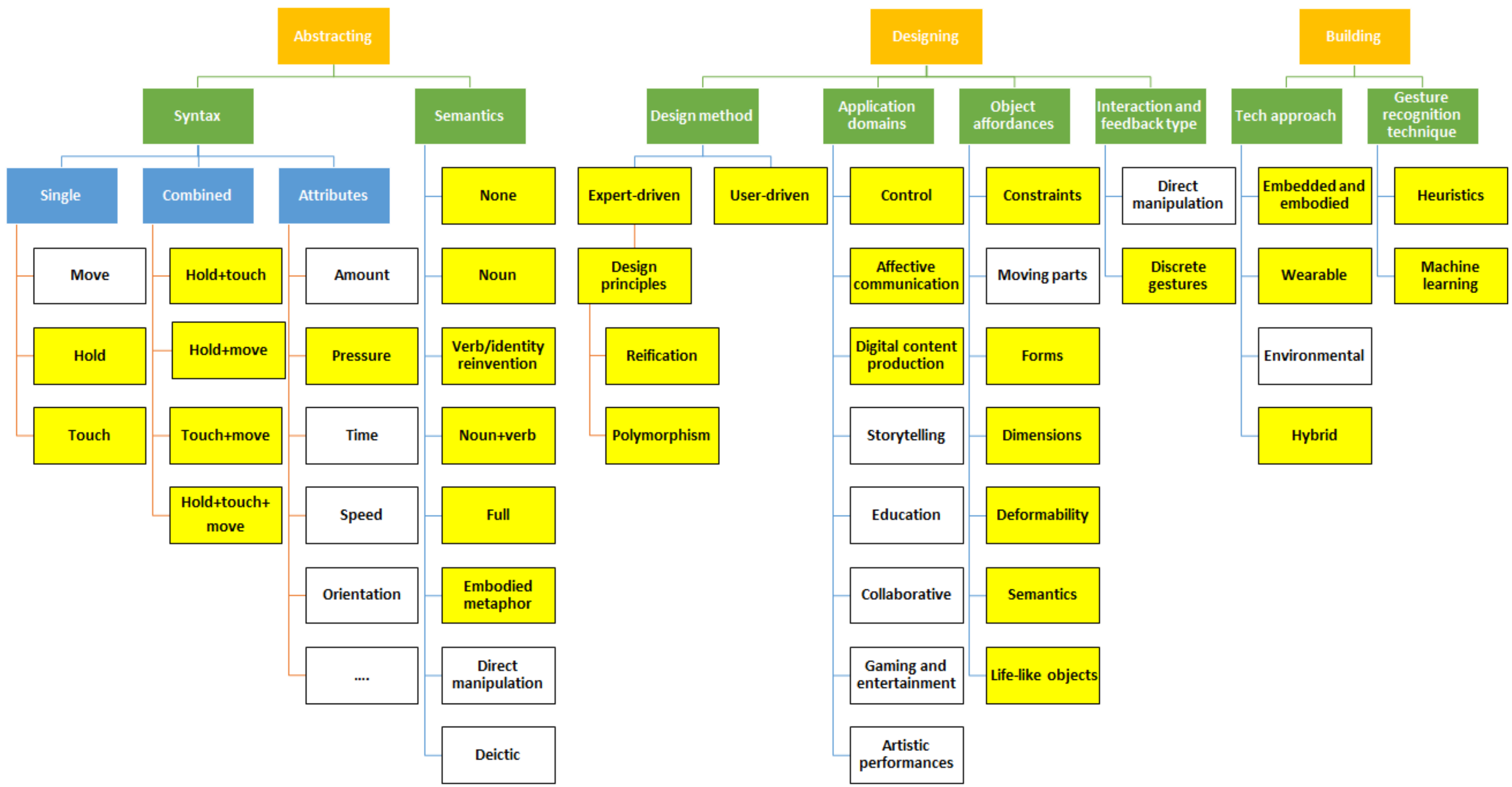

Figure A10. Elements of the TGIF design space explored in this thesis

Bibliography

1. Richard Alapack. 2007. Simulation in Cyberspace and Touch of the Flesh: Kissing, the Blush, the Hickey and the Caress. *Journal of Psychosocial Research in Cyberspace*, 1. 1-11.
2. Stanley H Ambrose. 2001. Paleolithic technology and human evolution. *Science*, 291 (5509). 1748-1753. http://dx.doi.org/10.1126/science.1059487
3. Roman Amstutz, Oliver Amft, Brian French, Asim Smailagic, Dan Siewiorek and Gerhard Troster. 2009. Performance analysis of an HMM-based gesture recognition using a wristwatch device. in *Computational Science and Engineering, 2009. CSE'09. International Conference on*, IEEE, 303-309. http://dx.doi.org/10.1109/CSE.2009.58
4. L Angelini, M Caon, D Lalanne, O Abou Khaled and E Mugellini. 2014. Une Lampe Anthropomorphique pour la Communication des Emotions. in *Actes d'IHM'14*, 207-212. http://dx.doi.org/10.1145/2670444.2670472
5. Leonardo Angelini, Jürgen Baumgartner, Francesco Carrino, Stefano Carrino, Maurizio Caon, Omar Khaled, Jürgen Sauer, Denis Lalanne, Elena Mugellini and Andreas Sonderegger. 2016. Comparing Gesture, Speech and Touch Interaction Modalities for In-Vehicle Infotainment Systems. in *Actes de la 28ième conférence francophone sur l'Interaction Homme-Machine*, ACM, 188-196. http://dx.doi.org/10.1145/3004107.3004118
6. Leonardo Angelini, Jürgen Baumgartner, Francesco Carrino, Stefano Carrino, Omar Abou Khaled, Jürgen Sauer, Denis Lalanne, Andreas Sonderegger and Elena Mugellini. 2016. Comparing Gesture, Speech and Touch Interaction Modalities for In-Vehicle Infotainment Systems. in *IHM'16*, To Appear. http://dx.doi.org/10.1145/3004107.3004118
7. Leonardo Angelini, Maurizio Caon, Francesco Carrino, Stefano Carrino, Denis Lalanne, Omar Abou Khaled and Elena Mugellini. 2013. WheelSense: Enabling Tangible Gestures on the Steering Wheel for In-Car Natural Interaction. in *Proc. HCII'13*, Las Vegas, NV, USA, 21-26 July 531-540. http://dx.doi.org/10.1007/978-3-642-39262-7_60
8. Leonardo Angelini, Maurizio Caon, Denis Lalanne, Omar Abou khaled and Elena Mugellini. 2015. Towards an Anthropomorphic Lamp for Affective Interaction. in *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '14*, Stanford, CA, USA, 16-19 January, 661-666. http://dx.doi.org/10.1145/2677199.2687914
9. Leonardo Angelini, Francesco Carrino, Stefano Carrino, Maurizio Caon, Omar Abou Khaled, Jürgen Baumgartner, Andreas Sonderegger, Denis Lalanne and Elena Mugellini. 2014. Gesturing on the Steering Wheel: a User-Elicited Taxonomy. in *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular*

Applications - AutomotiveUI '14, Seattle, WA, USA, 17-19 September, 1-8. http://dx.doi.org/10.1145/2667317.2667414
10. Leonardo Angelini, Francesco Carrino, Stefano Carrino, Maurizio Caon, Denis Lalanne, Omar Abou Khaled and Elena Mugellini. 2013. Opportunistic synergy: a classifier fusion engine for micro-gesture recognition. in *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13*, Eindhoven, Netherlands, 28-30 October, 30-37. http://dx.doi.org/10.1145/2516540.2516563
11. Leonardo Angelini, Omar Abou Khaled, Maurizio Caon, Elena Mugellini and Denis Lalanne. 2014. Hugginess: encouraging interpersonal touch through smart clothes. in *Proceedings of the 2014 ACM International Symposium on Wearable Computers Adjunct Program - ISWC '14 Adjunct*, Seattle, WA, USA, 13-17 September, 155-162. http://dx.doi.org/10.1145/2641248.2641356
12. Leonardo Angelini, Denis Lalanne, Elise van den Hoven, Omar Abou Khaled and Elena Mugellini. 2015. Move, Hold and Touch: A Framework for Tangible Gesture Interactive Systems. *Machines*, 3 (3). 173-207. http://dx.doi.org/10.3390/machines3030173
13. Leonardo Angelini, Denis Lalanne, Elise van den Hoven, Ali Mazalek, Omar Abou Khaled and Elena Mugellini. 2015. Tangible Meets Gestural: Comparing and Blending Post-WIMP Interaction Paradigms. in *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, ACM, 473-476. http://dx.doi.org/10.1145/2677199.2683583
14. Daniel Ashbrook, Kent Lyons and Thad Starner. 2008. An investigation into round touchscreen wristwatch interaction. in *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services - MobileHCI '08*, Amsterdam, The Netherlands, 2-5 September, ACM Press, 311-314. http://dx.doi.org/10.1145/1409240.1409276
15. Ayman Atia, Shin Takahashi and Jiro Tanaka. 2010. Smart Gesture Sticker: Smart Hand Gestures Profiles for Daily Objects Interaction. in *2010 IEEE/ACIS 9th International Conference on Computer and Information Science*, Yamagata, Japan, 18-20 August, Ieee, 482-487. http://dx.doi.org/10.1109/ICIS.2010.103
16. Luigi Atzori, Antonio Iera and Giacomo Morabito. 2010. The internet of things: A survey. *Computer networks*, 54 (15). 2787-2805. http://dx.doi.org/10.1016/j.comnet.2010.05.010
17. Kenneth Majlund Bach, Mads Gregers Jæger, Mikael B Skov and Nils Gram Thomassen. 2009. Interacting with in-vehicle systems: understanding, measuring, and evaluating attention. in *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*, British Computer Society, 453-462.
18. Kenneth Majlund Bach, Mads Gregers Jæger, Mikael B Skov and Nils Gram Thomassen. 2008. You can touch, but you can't look: interacting with in-vehicle systems. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1139-1148. http://dx.doi.org/10.1145/1357054.1357233
19. Saskia Bakker, Alissa N Antle and Elise van den Hoven. 2011. Embodied metaphors in tangible interaction design. *Personal and Ubiquitous Computing*, 16. 433-449. http://dx.doi.org/10.1007/s00779-011-0410-4
20. Massimo Banzi and Michael Shiloh. 2014. *Getting Started with Arduino: The Open Source Electronics Prototyping Platform*. Maker Media, Inc.
21. Thomas Baudel and Michel Beaudouin-Lafon. 1993. Charade: remote control of objects using free-hand gestures. *Commun. ACM*, 36 (7). 28-35. http://dx.doi.org/10.1145/159544.159562
22. Ayah Bdeir. 2009. Electronics as material: littleBits. in *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, ACM, 397-400. http://dx.doi.org/10.1145/1517664.1517743

23. Michel Beaudouin-Lafon. 2004. Designing interaction, not interfaces. in *Proceedings of the working conference on Advanced visual interfaces*, Gallipoli, Italy, 25-28 May, ACM, 15-22. http://dx.doi.org/10.1145/989863.989865
24. Michel Beaudouin-Lafon and Wendy E. Mackay. 2000. Reification, polymorphism and reuse: three principles for designing visual interfaces. in *Proceedings of the working conference on Advanced visual interfaces*, Palermo, Italy, 24 - 26 May, ACM, 102-109. http://dx.doi.org/10.1145/345513.345267
25. Victoria Bellotti, Maribeth Back, W Keith Edwards, Rebecca E Grinter, Austin Henderson and Cristina Lopes. 2002. Making sense of sensing systems: five questions for designers and researchers. in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 415-422. http://dx.doi.org/10.1145/503376.503450
26. Hrvoje Benko, Andrew D Wilson and Ravin Balakrishnan. 2008. Sphere: multi-touch interactions on a spherical display. in *Proceedings of the 21st annual ACM symposium on User interface software and technology*, ACM, 77-86. http://dx.doi.org/10.1145/1449715.1449729
27. Xiaojun Bi, Tomer Moscovich, Gonzalo Ramos, Ravin Balakrishnan and Ken Hinckley. 2008. An exploration of pen rolling for pen-based interaction. in *Proceedings of the 21st annual ACM symposium on User interface software and technology - UIST '08*, Monterey, California, USA, 19-22 October, ACM Press, 191-200. http://dx.doi.org/10.1145/1449715.1449745
28. D Bishop. 1992. Marble answering machine *Royal College of Art, Interaction Design*.
29. Richard A Bolt. 1980. "Put-that-there": Voice and gesture at the graphics interface. in *SIGGRAPH '80: Proceedings of the 7th annual conference on Computer graphics and interactive techniques*, Seattle, Washington, USA, 14-18 July 262-270. http://dx.doi.org/http://doi.acm.org/10.1145/800250.807503
30. Scott Brave and Andrew Dahley. 1997. inTouch: a medium for haptic interpersonal communication. in *CHI'97 Extended Abstracts on Human Factors in Computing Systems*, ACM, 363-364. http://dx.doi.org/10.1145/1120212.1120435
31. John Brooke. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry*, 189 (194). 4-7.
32. Leah Buechley, Mike Eisenberg, Jaime Catchen and Ali Crockett. 2008. The LilyPad Arduino: using computational textiles to investigate engagement, aesthetics, and diversity in computer science education. in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 423-432. http://dx.doi.org/10.1145/1357054.1357123
33. Karl Bühler. 1934. *Sprachtheorie*. Fischer, Oxford, England.
34. Laura Buttrick, Conor Linehan, Ben Kirman and Dan O'Hara. 2014. Fifty shades of CHI: the perverse and humiliating human-computer relationship. in *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, ACM, 825-834. http://dx.doi.org/10.1145/2559206.2578874
35. William Buxton. 1986. Chunking and phrasing and the design of human-computer dialogues. in *IFIP Congress*, 475-480.
36. Maurizio Caon, Leonardo Angelini, Omar Abou Khaled, Denis Lalanne, Yong Yue and Elena Mugellini. 2014. Affective interaction in smart environments. *Procedia Computer Science*, 32. 1016-1021. http://dx.doi.org/10.1016/j.procs.2014.05.527
37. Maurizio Caon, Leonardo Angelini, Yong Yue, Omar Abou Khaled and Elena Mugellini. 2013. Context-aware multimodal sharing of emotions. in *International Conference on Human-Computer Interaction*, Springer, 19-28. http://dx.doi.org/10.1007/978-3-642-39342-6_3
38. Maurizio Caon, Omar Abou Khaled, Elena Mugellini, Denis Lalanne and Leonardo Angelini. 2013. Ubiquitous Interaction for Computer Mediated Communication of Emotions. in *ACII*, 717-718. http://dx.doi.org/10.1109/ACII.2013.130
39. Maurizio Caon, Yong Yue, Julien Tscherrig, Elena Mugellini and Omar Abou Khaled. 2011. Context-Aware 3D Gesture Interaction Based on Multiple Kinects. in *AMBIENT 2011, The*

First International Conference on Ambient Computing, Applications, Services and Technologies, 7-12.

40. Barbara Caputo, K Sim, F Furesjo and Alex Smola. 2002. Appearance-based object recognition using SVMs: which kernel should I use? in *Proc of NIPS workshop on Statistical methods for computational experiments in visual processing and computer vision, Whistler*.
41. Francesco Carrino, Stefano Carrino, Maurizio Caon, Leonardo Angelini, Omar Abou Khaled and Elena Mugellini. 2012. In-Vehicle Natural Interaction Based on Electromyography. in *AutomotiveUI '12 Adjunct*, Portsmouth, NH, USA, 17-19 October, 23-24.
42. Stefano Carrino, Maurizio Caon, Omar Abou Khaled, Rolf Ingold and Elena Mugellini. 2013. Functional gestures for human-environment interaction. in *Human-Computer Interaction. Interaction Modalities and Techniques*, Las Vegas, NV, USA, 21-26 July Springer, 167-176. http://dx.doi.org/10.1007/978-3-642-39330-3_18
43. Stefano Carrino, Elena Mugellini, Omar Abou Khaled and Rolf Ingold. 2011. ARAMIS: toward a hybrid approach for human-environment interaction. in *Human-Computer Interaction. Towards Mobile and Intelligent Interaction Environments*, Springer, 165-174.
44. Stefano Carrino, Alexandre Péclat, Elena Mugellini, Omar Abou Khaled and Rolf Ingold. 2011. Humans and smart environments: a novel multimodal interaction approach. in *Proceedings of the 13th international conference on multimodal interfaces*, ACM, 105-112. http://dx.doi.org/10.1145/2070481.2070501
45. Jongeun Cha, Mohamad Eid, Lara Rahal and Abdulmotaleb El Saddik. 2008. HugMe: An interpersonal haptic communication system. in *Haptic Audio visual Environments and Games, 2008. HAVE 2008. IEEE International Workshop on*, IEEE, 99-102. http://dx.doi.org/10.1109/HAVE.2008.4685306
46. Angela Chang, Ben Resner, Brad Koerner, XingChen Wang and Hiroshi Ishii. 2001. LumiTouch: an emotional communication device. in *CHI'01 extended abstracts on Human factors in computing systems*, ACM, 313-314. http://dx.doi.org/10.1145/634067.634252
47. Jingyuan Cheng, Gernot Bahle and Paul Lukowicz. 2012. A simple wristband based on capacitive sensors for recognition of complex hand motions. in *2012 IEEE Sensors*.
48. Mauricio Cirelli and Ricardo Nakamura. 2014. A Survey on Multi-touch Gesture Recognition and Multi-touch Frameworks. in *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, ACM, 35-44. http://dx.doi.org/10.1145/2669485.2669509
49. Orville Leverne Clubb. 2007. Human-to-Computer-to-Human Interactions (HCHI) of the communications revolution. *interactions*, 14 (2). 35-39. http://dx.doi.org/10.1145/1229863.1229883
50. CuteCircuit Company. 2014. HugShirt.
51. Enrico Costanza, Samuel A Inverso and Rebecca Allen. 2005. Toward subtle intimate interfaces for mobile devices using an EMG controller. in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 481-489. http://dx.doi.org/10.1145/1054972.1055039
52. Chi Tai Dang and Elisabeth André. 2013. TabletopCars - Interaction with Active Tangible Remote Controlled Cars. in *Proc.TEI'13*, Barcelona, Spain, 10-13 February, 33-40. http://dx.doi.org/10.1145/2460625.2460630
53. Ferdinand De Saussure. 2011. *Course in general linguistics*. Columbia University Press.
54. Clarisse Sieckenius DeSouza. 2005. *The semiotic engineering of human-computer interaction*.
55. Beate Ditzen, Inga D Neumann, Guy Bodenmann, Bernadette von Dawans, Rebecca A Turner, Ulrike Ehlert and Markus Heinrichs. 2007. Effects of different kinds of couple interaction on cortisol and heart rate responses to stress in women. *Psychoneuroendocrinology*, 32 (5). 565-574. http://dx.doi.org/10.1016/j.psyneuen.2007.03.011
56. Tom Djajadiningrat. 2016. Inherently Meaningful. in *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, ACM, 2-2. http://dx.doi.org/10.1145/2839462.2883590

57. Tom Djajadiningrat, Ben Matthews and Marcelle Stienstra. 2007. Easy doesn't do it: skill and expression in tangible aesthetics. *Personal and Ubiquitous Computing*, 11. 657-676. http://dx.doi.org/10.1007/s00779-006-0137-9
58. Tom Djajadiningrat, Stephan Wensveen, Joep Frens and Kees Overbeeke. 2004. Tangible products: redressing the balance between appearance and action. *Personal and Ubiquitous Computing*, 8 (5). 294-309. http://dx.doi.org/10.1007/s00779-004-0293-8
59. Tanja Döring, Dagmar Kern, Paul Marshall, Max Pfeiffer, Johannes Schöning, Volker Gruhn and Albrecht Schmidt. 2011. Gestural interaction on the steering wheel: reducing the visual demand. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 483-492. http://dx.doi.org/10.1145/1978942.1979010
60. Paul Dourish. 2004. *Where The Action Is*. MIT Press, Cambridge, MA, USA.
61. Christoph Endres, Tim Schwartz and Christian A Müller. 2011. Geremin: 2D microgestures for drivers based on electric field sensing. in *Proceedings of the 16th international conference on Intelligent user interfaces*, ACM, 327-330. http://dx.doi.org/10.1145/1943403.1943457
62. Thomas Feix. Human Grasping Database.
63. A. Feldman, E.M. Tapia, S. Sadi, P. Maes and C. Schmandt. 2005. ReachMedia: On-the-move interaction with everyday objects. in *Ninth IEEE International Symposium on Wearable Computers (ISWC'05)*, Osaka, Japan, 18-21 October, IEEE, 52-59. http://dx.doi.org/10.1109/ISWC.2005.44
64. Alois Ferscha, Stefan Resmerita, Clemens Holzmann and Martin Reicho. 2005. Orientation sensing for gesture-based interaction with smart artifacts. *Computer Communications*, 28. 1552-1563. http://dx.doi.org/10.1016/j.comcom.2004.12.046
65. Tiffany Field. 2010. Touch for socioemotional and physical well-being: A review. *Developmental Review*, 30 (4). 367-383. http://dx.doi.org/10.1016/j.dr.2011.01.001
66. Kenneth P Fishkin, Anuj Gujar, Carlos Mochon and Roy Want. 1998. Squeeze Me , Hold Me , Tilt Me ! An Exploration of Manipulative User Interfaces. in *Proc. CHI'98*, Los Angeles, CA, USA, 18-23 April, 17-24. http://dx.doi.org/10.1145/274644.274647
67. Kenneth P Fishkin, Matthai Philipose and Adam D Rea. 2005. Hands-On RFID: Wireless Wearables for Detecting Use of Objects. in *ISWC*, 38-43. http://dx.doi.org/10.1109/ISWC.2005.25
68. KennethP. Fishkin. 2004. A taxonomy for and analysis of tangible interfaces. *Personal and Ubiquitous Computing*, 8. 347-358. http://dx.doi.org/10.1007/s00779-004-0297-4
69. GW Fitzmaurice. 1996. Graspable user interfaces *Citeseer*, University of Toronto, 167.
70. Lena M Forsell and Jan A Åström. 2012. Meanings of hugging: From greeting behavior to touching implications. *Comprehensive Psychology*, 1. 02.17. 21. CP. 01.13. http://dx.doi.org/10.2466/02.17.21.CP.1.13
71. Alberto Gallace and Charles Spence. 2010. The science of interpersonal touch: an overview. *Neuroscience & Biobehavioral Reviews*, 34 (2). 246-259. http://dx.doi.org/10.1016/j.neubiorev.2008.10.004
72. Maurizio Gentilucci and Michael C Corballis. 2006. From manual gesture to speech: A gradual transition. *Neuroscience & Biobehavioral Reviews*, 30 (7). 949-960. http://dx.doi.org/10.1016/j.neubiorev.2006.02.004
73. Luc Geurts and Vero Vanden Abeele. 2012. Splash Controllers : Game Controllers Involving the Uncareful Manipulation of Water. in *TEI'12*, Kingston, Ontario, Canada, 19-22 February, 183-186. http://dx.doi.org/10.1145/2148131.2148170
74. Arthur M Glenberg and David A Robertson. 1999. Indexical understanding of instructions. *Discourse Processes*, 28 (1). 1-26. http://dx.doi.org/10.1080/01638539909545067
75. J.W. Glos and Justine Cassell. 1997. Rosebud: Technological toys for storytelling. in *CHI'97 extended abstracts on Human factors in computing systems: looking to the future*, Atlanta, Georgia, USA, 22-27 March, 359-360. http://dx.doi.org/10.1145/1120212.1120433

76. Rolf Inge Godøy and Marc Leman. 2010. *Musical gestures: Sound, movement, and meaning*. Routledge.
77. Susan Goldin-Meadow, Carolyn Mylander, Jill de Villiers, Elizabeth Bates and Virginia Volterra. 1984. Gestural communication in deaf children: The effects and noneffects of parental input on early language development. *Monographs of the Society for Research in Child Development*, 49 (3-4). 1-151. http://dx.doi.org/10.2307/1165838
78. Ivan Golod, Felix Heidrich, Christian Möllering and Martina Ziefle. 2013. Design principles of hand gesture interfaces for microinteractions. in *Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces - DPPI '13*, Newcastle upon Tyne, UK, 3-5 September, 11-20. http://dx.doi.org/10.1145/2513506.2513508
79. Iván E González, Jacob O Wobbrock, Duen Horng Chau, Andrew Faulring and Brad A Myers. 2007. Eyes on the road, hands on the wheel: thumb-based interaction techniques for input on steering wheels. in *Proceedings of Graphics Interface 2007*, ACM, 95-102. http://dx.doi.org/10.1145/1268517.1268535
80. Saul Greenberg and Chester Fitchett. 2001. Phidgets: easy development of physical interfaces through physical widgets. in *Proceedings of the 14th annual ACM symposium on User interface software and technology*, ACM, 209-218. http://dx.doi.org/10.1145/502348.502388
81. Karen M Grewen, Bobbi J Anderson, Susan S Girdler and Kathleen C Light. 2003. Warm partner contact is related to lower cardiovascular reactivity. *Behavioral medicine*, 29 (3). 123-130. http://dx.doi.org/10.1080/08964280309596065
82. Antal Haans and Wijnand IJsselsteijn. 2006. Mediated social touch: a review of current research and future directions. *Virtual Reality*, 9 (2-3). 149-159. http://dx.doi.org/10.1007/s10055-005-0014-2
83. K.S. Kelly S Hale and Kay M K.M. Stanney. 2004. Deriving haptic design guidelines from human physiological, psychophysical, and neurological foundations. *Computer Graphics and Applications, IEEE*, 24 (2). 33-39. http://dx.doi.org/10.1109/MCG.2004.1274059
84. Edward T Hall, Ray L Birdwhistell, Bernhard Bock, Paul Bohannan, A Richard Diebold Jr, Marshall Durbin, Munro S Edmonson, JL Fischer, Dell Hymes and Solon T Kimball. 1968. Proxemics [and comments and replies]. *Current anthropology*. 83-108.
85. Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, 52. 139-183. http://dx.doi.org/10.1016/S0166-4115(08)62386-9
86. Bjoern Hartmann, Scott R Klemmer, Michael Bernstein and Nirav Mehta. 2005. d. tools: Visually prototyping physical UIs through statecharts. in *in Extended Abstracts of UIST 2005*, Citeseer.
87. Fabian Hemmert, Ulrike Gollner, Matthias Löwe, Anne Wohlauf and Gesche Joost. 2011. Intimate mobiles: grasping, kissing and whispering as a means of telecommunication in mobile phones. in *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, ACM, 21-24. http://dx.doi.org/10.1145/2037373.2037377
88. Ken Hinckley, Randy Pausch, John C Goble and Neal F Kassell. 1994. Passive real-world interface props for neurosurgical visualization. in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 452-458. http://dx.doi.org/10.1145/191666.191821
89. Lynn Hoff, Eva Hornecker and Sven Bertel. 2016. Modifying Gesture Elicitation: Do Kinaesthetic Priming and Increased Production Reduce Legacy Bias? in *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, ACM, 86-91. http://dx.doi.org/10.1145/2839462.2839472
90. Lars Erik Holmquist, Johan Redström and Peter Ljungstrand. 1999. Token-based access to digital information. in *Handheld and Ubiquitous Computing*, Springer, 234-245. http://dx.doi.org/10.1007/3-540-48157-5_22

91. Andreas Holzinger. 2002. Finger instead of mouse: touch screens as a means of enhancing universal access. in *Universal access theoretical perspectives, practice, and experience*, Springer, 387-397.

92. Eva Hornecker. 2012. Beyond affordance: tangibles' hybrid nature. in *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, Kingston, Ontario, Canada, 19 – 22 February, ACM, 175-182. http://dx.doi.org/10.1145/2148131.2148168

93. Eva Hornecker. 2010. Creative idea exploration within the structure of a guiding framework: the card brainstorming game. in *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*, ACM, 101-108. http://dx.doi.org/10.1145/1709886.1709905

94. Eva Hornecker. 2015. Tangible interaction. in (EDS.), M.S.A.R.F.D. ed. *The Glossary of Human Computer Interaction*.

95. Eva Hornecker and Jacob Buur. 2006. Getting a grip on tangible interaction: a framework on physical space and social interaction. in *Proceedings of the SIGCHI conference on Human Factors in computing systems*, Montréal, Québec, Canada, 22-27 April, 437-446. http://dx.doi.org/10.1145/1124772.1124838

96. Jörn Hurtienne and Johann Habakuk Israel. 2007. Image schemas and their metaphorical extensions: intuitive patterns for tangible interaction. in *Proceedings of the 1st international conference on Tangible and embedded interaction*, ACM, 127-134. http://dx.doi.org/10.1145/1226969.1226996

97. Hiroshi Ishii. 2008. Tangible bits: beyond pixels. in *Proceedings of the 2nd international conference on Tangible and embedded interaction*, ACM, xv-xxv. http://dx.doi.org/10.1145/1347390.1347392

98. Hiroshi Ishii, Dávid Lakatos, Leonardo Bonanni and Jean-Baptiste Jb Labrune. 2012. Radical Atoms : Beyond Tangible Bits , Toward Transformable Materials. *Interactions*, XIX. 38-51. http://dx.doi.org/10.1145/2065327.2065337

99. Robert J K Jacob, Audrey Girouard, Leanne M Hirshfield, Michael S Horn, Orit Shaer, Erin T. Solovey and Jamie Zigelbaum. 2008. Reality-Based Interaction : A Framework for Post-WIMP Interfaces. in *CHI'08*, Florence, Italy, 5-10 April, 201-210. http://dx.doi.org/10.1145/1357054.1357089

100. Hans-Christian Jetter, Harald Reiterer and Florian Geyer. 2014. Blended Interaction: understanding natural human–computer interaction in post-WIMP interactive spaces. *Personal and Ubiquitous Computing*, 18 (5). 1139-1158. http://dx.doi.org/10.1007/s00779-013-0725-4

101. Sergi Jordà, Günter Geiger, Marcos Alonso and Martin Kaltenbrunner. 2007. The reacTable : Exploring the Synergy between Live Music Performance and Tabletop Tangible Interfaces. in *TEI '07 Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, Baton Rouge, Louisiana, USA, 15-17 February, 139-146. http://dx.doi.org/10.1145/1226969.1226998

102. Mark Joselli and Esteban Clua. 2009. grmobile: A framework for touch and accelerometer gesture recognition for mobile games. in *Games and Digital Entertainment (SBGAMES), 2009 VIII Brazilian Symposium on*, IEEE, 141-150. http://dx.doi.org/10.1109/SBGAMES.2009.24

103. Département Fédéral de Police et Justice. 1977. Teaching manual for the formation and examination of driving instructors (Manuel d'enseignement pour la formation et l'examen des moniteurs de conduite).

104. Martin Kaltenbrunner and Ross Bencina. 2007. reacTIVision: a computer-vision framework for table-based tangible interaction. in *Proceedings of the 1st international conference on Tangible and embedded interaction*, ACM, 69-74. http://dx.doi.org/10.1145/1226969.1226983

105. Martin Kaltenbrunner, Till Bovermann, Ross Bencina and Enrico Costanza. 2005. TUIO: A protocol for table-top tangible user interfaces. in *Proc. of the The 6th Int'l Workshop on Gesture in Human-Computer Interaction and Simulation*, 1-5.

106. Dietrich Kammer, Mandy Keck, Georg Freitag and Markus Wacker. 2010. Taxonomy and overview of multi-touch frameworks: Architecture, scope and features. in *Workshop on Engineering Patterns for Multitouch Interfaces*.

107. Dietrich Kammer, Jan Wojdziak, Mandy Keck, Rainer Groh and Severin Taranko. 2010. Towards a formalization of multi-touch gestures. in *ACM International Conference on Interactive Tabletops and Surfaces - ITS '10*, Saarbrücken, Germany, 7-10 November, ACM Press, 49-58. http://dx.doi.org/10.1145/1936652.1936662

108. Maria Karam. 2006. A framework for research and design of gesture-based human computer interactions *Science*, 198.

109. Lynn E Kauffman. 1971. Tacesics, the study of touch: a model for proxemic analysis. *Semiotica*, 4 (2). 149-161. http://dx.doi.org/10.1515/semi.1971.4.2.149

110. Kathleen Keating. 1994. *The hug therapy book*. Hazelden Publishing.

111. John F Kelley. 1984. An iterative design methodology for user-friendly natural language office information applications. *ACM Transactions on Information Systems (TOIS)*, 2 (1). 26-41. http://dx.doi.org/10.1145/357417.357420

112. Adam Kendon. 1996. An agenda for gesture studies. *Semiotic review of books*, 7 (3). 8-12.

113. Adam Kendon. 1986. Current issues in the study of gesture. in *The biological foundations of gestures: Motor and semiotic aspects*, 23-47.

114. Hamed Ketabdar, Peyman Moghadam and Mehran Roshandel. 2012. Pingu: A new miniature wearable device for ubiquitous computing environments. in *Proceedings - 2012 6th International Conference on Complex, Intelligent, and Software Intensive Systems, CISIS 2012*, Palermo, Italy, 4-6 July, 502-506. http://dx.doi.org/10.1109/CISIS.2012.123

115. Floris Kimman, Hans Weda, E. van den Hoven, T. de Zeeuw and Steven Luitjens. 2011. Spinning in control: design exploration of a cross-device remote. in *TEI'11*, Funchal, Portugal, 22-26 January, ACM, 189-192. http://dx.doi.org/10.1145/1935701.1935738

116. Kenrick Kin, Björn Hartmann, Tony DeRose and Maneesh Agrawala. 2012. Proton: multitouch gestures as regular expressions. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2885-2894. http://dx.doi.org/10.1145/2207676.2208694

117. Kenrick Kin, Björn Hartmann, Tony DeRose and Maneesh Agrawala. 2012. Proton++: a customizable declarative multitouch framework. in *Proceedings of the 25th annual ACM symposium on User interface software and technology*, ACM, 477-486. http://dx.doi.org/10.1145/2380116.2380176

118. F Klompmaker, Karsten Nebe and Alex Fast. 2012. dSensingNI: a framework for advanced tangible interaction using a depth camera. in *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, Kingston, Ontario, Canada, 19-22 February, 217-224. http://dx.doi.org/10.1145/2148131.2148179

119. Boriana Koleva, Steve Benford, Kher Hui Ng and Tom Rodden. 2003. A Framework for Tangible User Interfaces. in *Physical Interaction (PI03) - Workshop on RealWorld User Interfaces*, Udine, Italy, 8 September, 46-50.

120. Robert Kowalski, Sebastian Loehmann and Doris Hausen. 2013. cubble: a multi-device hybrid approach supporting communication in long-distance relationships. in *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, ACM, 201-204. http://dx.doi.org/10.1145/2460625.2460656

121. Shunsuke Koyama, Yuta Sugiura, Masa Ogata, Anusha Withana, Yuji Uema, Makoto Honda, Sayaka Yoshizu, Chihiro Sannomiya, Kazunari Nawa and Masahiko Inami. 2014. Multi-touch steering wheel for in-car tertiary applications using infrared sensors. in *Proceedings of the 5th*

Augmented Human International Conference, ACM, 5. http://dx.doi.org/10.1145/2582051.2582056

122. Sébastien Kubicki, Sophie Lepreux and Christophe Kolski. 2012. RFID-driven situation awareness on TangiSense, a table interacting with tangible objects. *Personal and Ubiquitous Computing*, 16 (8). 1079-1094. http://dx.doi.org/10.1007/s00779-011-0442-9
123. Christine Kühnel, Tilo Westermann, Fabian Hemmert, Sven Kratz, Alexander Müller and Sebastian Möller. 2011. I'm home: Defining and evaluating a gesture set for smart-home control. *International Journal of Human-Computer Studies*, 69 (11). 693-704. http://dx.doi.org/10.1016/j.ijhcs.2011.04.005
124. Gordon Kurtenbach and Eric A Hulteen. 1990. Gestures in human-computer communication. in *The art of human-computer interface design*, 309-317.
125. Byron Lahey, Audrey Girouard, Winslow Burleson and Roel Vertegaal. 2011. PaperPhone: Understanding the Use of Bend Gestures in Mobile Devices with Flexible Electronic Paper Displays. in *Proc. CHI*, Vancouver, BC, Canada, 7-12 May, 1303-1312. http://dx.doi.org/10.1145/1978942.1979136
126. Angelini Leonardo, Abou Khaled Omar, Mugellini Elena and Lalane Denis. 2014. Les qualités tangibles de la lampe anthropomorphique interactive. in *IHM'14, 26e conférence francophone sur l'Interaction Homme-Machine*, 92-93.
127. James R Lewis and Jeff Sauro. 2009. The factor structure of the system usability scale. in *International Conference on Human Centered Design*, Springer, 94-103. http://dx.doi.org/10.1007/978-3-642-02806-9_12
128. Zhe Li, Sven Wachsmuth, Jannik Fritsch and Gerhard Sagerer. 2007. View-adaptive manipulative action recognition for robot companions. in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, San Diego, CA, USA, 29 October - 2 November, IEEE, 1028-1033. http://dx.doi.org/10.1109/IROS.2007.4399224
129. Kathleen C Light, Karen M Grewen and Janet A Amico. 2005. More frequent partner hugs and higher oxytocin levels are linked to lower blood pressure and heart rate in premenopausal women. *Biological psychology*, 69 (1). 5-21. http://dx.doi.org/10.1016/j.biopsycho.2004.11.002
130. Jeong-Mook Lim, Dong-Woo Lee, Bae-Sun Kim, Il-Yeon Cho and Jae-Cheol Ryou. 2010. Recognizing hand gestures using wrist shapes. in *2010 Digest of Technical Papers International Conference on Consumer Electronics (ICCE)*, IEEE, 197-198. http://dx.doi.org/10.1109/ICCE.2010.5418904
131. Roman Lissermann. 2012. FoldMe : Interacting with Double-sided Foldable Displays. in *Proc.TEI'12*, Kingston, Ontario, Canada, 19-22 February, 33-40. http://dx.doi.org/10.1145/2148131.2148142
132. Jiayang Liu, Lin Zhong, Jehan Wickramasuriya and Venu Vasudevan. 2009. uWave: Accelerometer-based personalized gesture recognition and its applications. *Pervasive and Mobile Computing*, 5 (6). 657-675. http://dx.doi.org/10.1016/j.pmcj.2009.07.007
133. Pedro Lopes, Patrik Jonell and Patrick Baudisch. 2015. Affordance++: Allowing Objects to Communicate Dynamic Use. in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, Seoul, Republic of Korea, 18-23 April, ACM, 2515-2524. http://dx.doi.org/10.1145/2702123.2702128
134. Einar Sneve Martinussen and Timo Arnall. 2009. Designing with RFID. in *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, Cambridge, United Kingdom, ACM, 343-350. http://dx.doi.org/10.1145/1517664.1517734
135. Stefan Mattes. 2003. The lane-change-task as a tool for driver distraction evaluation. *Quality of Work and Products in Enterprises of the Future*. 57-60.
136. Ben Matthews. 2006. Grammar, meaning and movement-based interaction. in *Proc. OZCHI '06*, Sydney, Australia, 20-24 November, ACM Press, 405-408. http://dx.doi.org/10.1145/1228175.1228258

137. Ali Mazalek and Elise van den Hoven. 2009. Framing tangible interaction frameworks. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 23. 225. http://dx.doi.org/10.1017/S0890060409000201
138. David McNeill. 1992. *Hand and mind: What gestures reveal about thought*. University of Chicago press.
139. George A Miller. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, 63 (2). 81. http://dx.doi.org/10.1037/0033-295X.101.2.343
140. Eve Mitsopoulos-Rubens, Margaret J Trotter and Michael G Lenné. 2011. Effects on driving performance of interacting with an in-vehicle music player: A comparison of three interface layout concepts for information presentation. *Applied ergonomics*, 42 (4). 583-591. http://dx.doi.org/10.1016/j.apergo.2010.08.017
141. Ashley Montagu. 1971. *Touching: The human significance of the skin*.
142. Elisa Morganti, Leonardo Angelini, Andrea Adami, Denis Lalanne, Leandro Lorenzelli and Elena Mugellini. 2012. A Smart Watch with Embedded Sensors to Recognize Objects, Grasps and Forearm Gestures. in *Procedia Engineering, International Symposium on Robotics and Intelligent Sensors 2012*, Kuching, Sarawak, Malaysia, 4-6 September, 1169-1175. http://dx.doi.org/10.1016/j.proeng.2012.07.297
143. Meredith Ringel Morris, Andreea Danielescu, Steven Drucker, Danyel Fisher, Bongshin Lee, m. c. schraefel and Jacob O. Wobbrock. 2014. Reducing legacy bias in gesture elicitation studies. *interactions*, 21 (3). 40-45. http://dx.doi.org/10.1145/2591689
144. Meredith Ringel Morris, Jacob O Wobbrock and Andrew D Wilson. 2010. Understanding users' preferences for surface gestures. in *Proceedings of graphics interface 2010*, Canadian Information Processing Society, 261-268.
145. Florian'Floyd' Mueller, Frank Vetere, Martin R Gibbs, Jesper Kjeldskov, Sonja Pedell and Steve Howard. 2005. Hug over a distance. in *CHI'05 extended abstracts on Human factors in computing systems*, ACM, 1673-1676. http://dx.doi.org/10.1145/1056808.1056994
146. Elena Mugellini, Elisa Rubegni, Sandro Gerardi and Omar Abou Khaled. 2007. Using personal objects as tangible interfaces for memory recollection and sharing. in *Proceedings of the 1st international conference on Tangible and embedded interaction*, Baton Rouge, Louisiana, 15-17 February, ACM, 231-238. http://dx.doi.org/10.1145/1226969.1227016
147. Donald A Norman. 2010. Natural user interfaces are not natural. *interactions*, 17 (3). 6-10. http://dx.doi.org/10.1145/1744161.1744163
148. Michael Nunes, Saul Greenberg and Carman Neustaedter. 2008. Sharing digital photographs in the home through physical mementos, souvenirs, and keepsakes. in *Proceedings of the 7th ACM conference on Designing interactive systems*, ACM, 250-260. http://dx.doi.org/10.1145/1394445.1394472
149. Rafael Nunes, Fabio Rito and Carlos Duarte. 2015. TACTIC: An API for Touch and Tangible Interaction. in *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, ACM, 125-132. http://dx.doi.org/10.1145/2677199.2680592
150. Dan O'Sullivan and Tom Igoe. 2004. *Physical computing: sensing and controlling the physical world with computers*. Course Technology Press.
151. Swiss Federation - Federal Statistical Office. 2015. Mobility and Transport - Pocket Statistics 2015.
152. Johanna Okerlund, Evan Segreto, Casey Grote, Lauren Westendorf, Anja Scholze, Romie Littrell and Orit Shaer. 2016. SynFlo: A Tangible Museum Exhibit for Exploring Bio-Design. in *Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, Eindhoven, Netherlands, ACM, 141-149. http://dx.doi.org/10.1145/2839462.2839488
153. Jerome Pasquero, S.J. Stobbe and Noel Stonehouse. 2011. A haptic wristwatch for eyes-free interactions. in *Proceedings of the 2011 annual conference on Human factors in computing*

systems, Vancouver, BC, Canada, 7–12 May ACM, 3257-3266. http://dx.doi.org/10.1145/1978942.1979425
154. Annie Pauzié. 2008. A method to assess the driver mental workload: The driving activity load index (DALI). *IET Intelligent Transport Systems*, 2 (4). 315-322. http://dx.doi.org/10.1049/iet-its:20080023
155. Vladimir I Pavlovic, Rajeev Sharma and Thomas S Huang. 1997. Visual interpretation of hand gestures for human-computer interaction: A review. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19 (7). 677-695. http://dx.doi.org/Doi 10.1109/34.598226
156. Charles Sanders Peirce. 1902. *Logic as semiotic: The theory of signs*.
157. Simob Perrault, Eric Lecolinet, James Eagan and Yves Guiard. 2013. WatchIt : Simple Gestures and Eyes-free Interaction for Wristwatches and Bracelets. in *CHI'13*, Paris, France, 27 April - 2 May, 1-10. http://dx.doi.org/10.1145/2470654.2466192
158. L Peters. 2006. Nabaztag wireless communicator. *Personal Computer World*, 2.
159. Daniela Petrelli, Alessandro Soranzo, Luigina Ciolfi and John Reidy. 2016. Exploring the aesthetics of tangible interaction: experiments on the perception of hybrid objects. in *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction*, ACM, 100-108. http://dx.doi.org/10.1145/2839462.2839478
160. Bastian Pfleging, Stefan Schneegass and Albrecht Schmidt. 2012. Multimodal interaction in the car: combining speech and gestures on the steering wheel. in *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ACM, 155-162. http://dx.doi.org/10.1145/2390256.2390282
161. Robert Plutchik. 2001. The Nature of Emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist*, 89 (4). 344-350. http://dx.doi.org/10.1511/2001.4.344
162. Sara Price and Yvonne Rogers. 2004. Let's get physical: The learning benefits of interacting in digitally augmented physical spaces. *Computers & Education*, 43. 137-151. http://dx.doi.org/10.1016/j.compedu.2003.12.009
163. Francis KH Quek. 1994. Toward a vision-based hand gesture interface. in *Virtual Reality Software and Technology Conference*, 17-29.
164. Francis Quek, David McNeill, Robert Bryll, Susan Duncan, Xin-Feng Ma, Cemil Kirbas, Karl E McCullough and Rashid Ansari. 2002. Multimodal human discourse: gesture and speech. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 9 (3). 171-193. http://dx.doi.org/10.1145/568513.568514
165. Francis Quek, David Mcneill, Robert Bryll, Susan Duncan, Xin-Feng Ma, Cemil Kirbas, Karl E. Mccullough and Rashid Ansari. 2002. Multimodal human discourse: gesture and speech. *ACM Transactions on Computer-Human Interaction*, 9. 171-193. http://dx.doi.org/10.1145/568513.568514
166. Luis Radford. 2008. Why do gestures matter? Sensuous cognition and the palpability of mathematical meanings. *Educational Studies in Mathematics*, 70. 111-126. http://dx.doi.org/10.1007/s10649-008-9127-3
167. Giuseppe Raffa, Jinwon Lee, Lama Nachman and Junehwa Song. 2010. Don't slow me down: Bringing energy efficiency to continuous gesture recognition. in *International Symposium on Wearable Computers (ISWC) 2010*, IEEE, 1-8. http://dx.doi.org/10.1109/ISWC.2010.5665872
168. ASM Rahman, Jamal Saboune and Abdulmotaleb El Saddik. 2011. Motion-path based in car gesture control of the multimedia devices. in *Proceedings of the first ACM international symposium on Design and analysis of intelligent vehicular networks and applications*, ACM, 69-76. http://dx.doi.org/10.1145/2069000.2069013
169. Siddharth S Rautaray and Anupam Agrawal. 2015. Vision based hand gesture recognition for human computer interaction: a survey. *Artificial Intelligence Review*, 43 (1). 1-54. http://dx.doi.org/10.1007/s10462-012-9356-9

170. Jun Rekimoto. 2001. Gesturewrist and gesturepad: Unobtrusive wearable interaction devices. in *Wearable Computers, 2001. Proceedings. Fifth International Symposium on*, IEEE, 21-27. http://dx.doi.org/10.1109/ISWC.2001.962092
171. Patrick Reuter, Guillaume Riviere, Nadine Couture, Stephanie Mahut and Loic Espinasse. 2010. ArcheoTUI—Driving virtual reassemblies with tangible 3D interaction. *Journal on Computing and Cultural Heritage (JOCCH)*, 3 (2). 4. http://dx.doi.org/10.1145/1841317.1841319
172. Bradley J Rhodes, Nelson Minar and Joshua Weaver. 1999. Wearable computing meets ubiquitous computing: Reaping the best of both worlds. in *Wearable Computers, 1999. Digest of Papers. The Third International Symposium on*, IEEE, 141-149. http://dx.doi.org/10.1109/ISWC.1999.806695
173. Andreas Riener. 2012. Gestural interaction in vehicular applications. *Computer*, 45 (4). 42-47. http://dx.doi.org/10.1109/MC.2012.108
174. Andreas Riener, Alois Ferscha, Florian Bachmair, Patrick Hagmüller, A Lemme, Dominik Muttenthaler, David Pühringer, Harald Rogner, Adrian Tappe and Florian Weger. 2013. Standardization of the in-car gesture interaction space. in *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ACM, 14-21. http://dx.doi.org/10.1145/2516540.2516544
175. Andreas Riener and Philipp Wintersberger. 2011. Natural, intuitive finger based input as substitution for traditional vehicle control. in *Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ACM, 159-166. http://dx.doi.org/10.1145/2381416.2381442
176. Toni Robertson. 1997. Cooperative work and lived cognition: a taxonomy of embodied actions. in *Proceedings of the Fifth European Conference on Computer Supported Cooperative Work*, Springer, 205-220. http://dx.doi.org/10.1007/978-94-015-7372-6_14
177. Claire Rowland, Elizabeth Goodman, Martin Charlier, Ann Light and Alfred Lui. 2015. *Designing Connected Products: UX for the Consumer Internet of Things*. " O'Reilly Media, Inc.".
178. K. Ryokai, H. Raffle, H. Horii and Y. Mann. 2010. Tangible video bubbles. in *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*, Atlanta, Georgia, USA, 10-15 April, ACM, 2775-2784. http://dx.doi.org/10.1145/1753846.1753863
179. T Scott Saponas, Desney S Tan, Dan Morris, Ravin Balakrishnan, Jim Turner and James A Landay. 2009. Enabling always-available input with muscle-computer interfaces. in *Proceedings of the 22nd annual ACM symposium on User interface software and technology*, ACM, 167-176. http://dx.doi.org/10.1145/1622176.1622208
180. Munehiko Sato, Ivan Poupyrev and Chris Harrison. 2012. 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)*, Austin, Texas, USA, 5-10 May, 483-492. http://dx.doi.org/10.1145/2207676.2207743
181. Urs Schallberger. 2005. Kurzskalen zur Erfassung der Positiven Aktivierung, Negativen Aktivierung und Valenz in Experience Sampling Studien (PANAVA-KS). *Theoretische und methodische Grundlagen, Konstruktvalidität und psychometrische Eigenschaften bei der Beschreibung intra-und interindividueller Unterschiede*.
182. Emanuel A Schegloff. 1984. On some gestures' relation to talk. *Structures of social action*, 266295. http://dx.doi.org/10.1017/CBO9780511665868.018
183. Albrecht Schmidt, H-W Gellersen and Christian Merz. 2000. Enabling implicit human computer interaction: a wearable RFID-tag reader. in *Wearable Computers, The Fourth International Symposium on*, IEEE, 193-194. http://dx.doi.org/10.1109/ISWC.2000.888497

184. Michael Schmitz. 2011. Concepts for life-like interactive objects. in *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, Funchal, Portugal, 22-26 January, ACM Press, 157-164. http://dx.doi.org/10.1145/1935701.1935732

185. Orit Shaer and Eva Hornecker. 2010. Tangible user interfaces: past, present, and future directions. *Foundations and Trends in Human-Computer Interaction*, 3 (1–2). 1-137. http://dx.doi.org/10.1561/1100000026

186. Orit Shaer and Robert JK Jacob. 2009. A specification paradigm for the design and implementation of tangible user interfaces. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 16 (4). 20. http://dx.doi.org/10.1145/1614390.1614395

187. Orit Shaer, Nancy Leland, EduardoH. Calvillo-Gamez and RobertJ.K. Jacob. 2004. The TAC paradigm: specifying tangible user interfaces. *Personal and Ubiquitous Computing*, 8. 359-369. http://dx.doi.org/10.1007/s00779-004-0298-3

188. Ehud Sharlin, Benjamin Watson, Yoshifumi Kitamura, Fumio Kishino and Yuichi Itoh. 2004. On tangible user interfaces, humans and spatiality. *Personal and Ubiquitous Computing*, 8 (5). 338-346. http://dx.doi.org/10.1007/s00779-004-0296-5

189. Timothy M. Simon, Bruce H. Thomas, Ross T. Smith and Mark Smith. 2013. Adding input controls and sensors to RFID tags to support dynamic tangible user interfaces. in *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction*, Munich, Germany, ACM, 165-172. http://dx.doi.org/10.1145/2540930.2540979

190. Jonas Sjöström and Göran Goldkuhl. 2003. The semiotics of user interfaces–a socio-pragmatic perspective. in *6th International Workshop on Organisational Semiotics*, 11-12. http://dx.doi.org/10.1007/1-4020-2162-3_14

191. Andreas Sonderegger and Juergen Sauer. 2010. The influence of design aesthetics in usability testing: Effects on user performance and perceived usability. *Applied ergonomics*, 41 (3). 403-410. http://dx.doi.org/10.1016/j.apergo.2009.09.002

192. Hyunyoung Song, Hrvoje Benko, F. Guimbretiere, Shahram Izadi, Xiang Cao and Ken Hinckley. 2011. Grips and gestures on a multi-touch pen. in *Proceedings of the 2011 annual conference on Human factors in computing systems*, Vancouver, BC, Canada, 7-12 May, ACM, 1323-1332. http://dx.doi.org/10.1145/1978942.1979138

193. Iris Soute. 2013. Head Up Games: On the design, creation and evaluation of interactive outdoor games for children, Eindhoven University of Technology 232.

194. Lucio Davide Spano, Antonio Cisternino, Fabio Paternò and Gianni Fenu. 2013. GestIT: a declarative and compositional framework for multiplatform gesture definition. in *Proceedings of the 5th ACM SIGCHI symposium on Engineering interactive computing systems*, ACM, 187-196. http://dx.doi.org/10.1145/2494603.2480307

195. Thad Starner, Jake Auxier, Daniel Ashbrook and Maribeth Gandy. 2000. The gesture pendant: A self-illuminating, wearable, infrared computer vision system for home automation control and medical monitoring. in *Wearable computers, the fourth international symposium on*, IEEE, 87-94. http://dx.doi.org/10.1109/ISWC.2000.888469

196. Maja Stikic, Tâm Huynh, Kristof Van Laerhoven and Bernt Schiele. 2008. ADL recognition based on the combination of RFID and accelerometer sensing. in *2008 Second International Conference on Pervasive Computing Technologies for Healthcare*, IEEE, 258-263. http://dx.doi.org/10.1109/PCTHEALTH.2008.4571084

197. Rob Strong and Bill Gaver. 1996. Feather, scent and shaker: supporting simple intimacy. in *Proceedings of CSCW*, 29-30.

198. Joshua Tanenbaum, Karen Tanenbaum and Alissa Antle. 2010. The Reading Glove : Designing Interactions for Object-Based Tangible Storytelling. in *Proc.AH'10*, Megève, France, 2-3 April, 1-9. http://dx.doi.org/10.1145/1785455.1785474

199. Brandon Taylor and V. Michael Bove. 2009. Graspables : Grasp-Recognition as a User Interface. in *CHI'09*, Boston, MA, USA, 4-9 April, 917-925. http://dx.doi.org/10.1145/1518701.1518842

200. James Keng Soon Teh, Adrian David Cheok, Roshan L Peiris, Yongsoon Choi, Vuong Thuong and Sha Lai. 2008. Huggy Pajama: a mobile parent and child hugging communication system. in *Proceedings of the 7th international conference on Interaction design and children*, ACM, 250-257. http://dx.doi.org/10.1145/1463689.1463763

201. Telegraph. 2010. Britons spend more time driving than socialising. *http://www.telegraph.co.uk/motoring/news/8287098/Britons-spend-more-time-driving-than-socialising.html*

202. Eric Tobias, Valérie Maquil and Thibaud Latour. 2015. TULIP: a widget-based software framework for tangible tabletop interfaces. in *Proceedings of the 7th ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, ACM, 216-221. http://dx.doi.org/10.1145/2774225.2775080

203. Dzmitry Tsetserukou. 2010. HaptiHug: a novel haptic display for communication of hug over a distance. in *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, Springer, 340-347. http://dx.doi.org/10.1007/978-3-642-14064-8_49

204. Sherry Turkle. 2012. *Alone together: Why we expect more from technology and less from each other*. Basic books.

205. B. Ullmer and H. Ishii. 2000. Emerging frameworks for tangible user interfaces. *IBM Systems Journal*, 39. 915-931. http://dx.doi.org/10.1147/sj.393.0915

206. Brigg Ullmer and Albert Schmidt. 2007. Proceedings of the 1st international conference on Tangible and embedded interaction. in, Baton Rouge, Louisiana, ACM, 296.

207. John Underkoffler and Hiroshi Ishii. 1999. Urp: a luminous-tangible workbench for urban planning and design. in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, ACM, 386-393. http://dx.doi.org/10.1145/302979.303114

208. Frits Vaandrager. 2010. What is a Good Model? *http://www.cs.ru.nl/~fvaan/PV/what_is_a_good_model.html*

209. Consuelo Valdes, Diana Eastman, Casey Grote, Shantanu Thatte, Orit Shaer, Ali Mazalek, Brygg Ullmer and Miriam K. Konkel. 2014. Exploring the design space of gestural interaction with active tokens through user-defined gestures. in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, Toronto, Ontario, Canada, 26 April - 01 May, ACM, 4107-4116. http://dx.doi.org/10.1145/2556288.2557373

210. Myriam Valentini, Udo Kischka and Peter W Halligan. 2008. Residual haptic sensation following stroke using ipsilateral stimulation. *Journal of Neurology, Neurosurgery & Psychiatry*, 79 (3). 266-270. http://dx.doi.org/10.1136/jnnp.2007.120279

211. Albert van Breemen, Xue Yan and Bernt Meerbeek. 2005. iCat: an animated user-interface robot with personality. in *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*, ACM, 143-144. http://dx.doi.org/10.1145/1082473.1082823

212. Elise van den Hoven and Berry Eggen. 2004. Tangible Computing in Everyday Life: Extending Current Frameworks for Tangible User Interfaces with Personal Objects. *Ambient Intelligence*, 3295. 230-242. http://dx.doi.org/10.1007/978-3-540-30473-9_22

213. Elise van den Hoven and Ali Mazalek. 2011. Grasping gestures: Gesturing with physical artifacts. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 25. 255-271. http://dx.doi.org/10.1017/S0890060411000072

214. Elise Van Den Hoven, Evelien Van De Garde-Perik, Serge Offermans, Koen Van Boerdonk and Kars Michiel H Lenssen. 2013. Moving tangible interaction systems to the next level. *Computer*, 46. 70-76. http://dx.doi.org/10.1109/MC.2012.360

215. Jelle van Dijk, Camille Moussette, Stoffel Kuenen and Caroline Hummels. 2013. Radical clashes: what tangible interaction is made of. in *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, ACM, 323-326. http://dx.doi.org/10.1145/2460625.2460680

216. Jelle van Dijk, Remko van der Lugt and Caroline Hummels. 2014. Beyond distributed representation: embodied cognition design supporting socio-sensorimotor couplings. in *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction*, ACM, 181-188. http://dx.doi.org/10.1145/2540930.2540934
217. Cati Vaucelle and H. Ishii. 2008. Picture This! Film assembly using toy gestures. in *Proceedings of the 10th international conference on Ubiquitous computing*, Seoul, Korea, 21-24 September, Citeseer, 350-359. http://dx.doi.org/http://doi.acm.org/10.1145/1409635.1409683
218. Nicolas Villar and Steve Hodges. 2010. The peppermill: a human-powered user interface device. in *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*, ACM, 29-32. http://dx.doi.org/10.1145/1709886.1709893
219. Nicolas Villar, James Scott, Steve Hodges, Kerry Hammil and Colin Miller. 2012. . NET gadgeteer: a platform for custom devices. in *Pervasive Computing*, Springer, 216-233.
220. Chris Walker. 2012. *Getting started with Netduino*. " O'Reilly Media, Inc.".
221. Judith Weisberg and Maxine R Haberman. 1989. A therapeutic hugging week in a geriatric facility. *Journal of Gerontological Social Work*, 13 (3-4). 181-186. http://dx.doi.org/Doi 10.1300/J083v13n03_13
222. Mark Weiser. 1991. The computer for the 21st century. *Scientific american*, 265 (3). 94-104. http://dx.doi.org/10.1145/329124.329126
223. Pierre Wellner, Wendy Mackay and Rich Gold. 1993. Computer-Augmented Environments: Back to the Real World. *Communications of the ACM*, 36. 24-26. http://dx.doi.org/10.1145/159544.159555
224. S A G Wensveen, J P Djajadiningrat and C J Overbeeke. 2004. Interaction frogger: a design framework to couple action and function through feedback and feedforward. in *DIS '04 Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, Cambridge, Massachusetts, USA, 1-4 August, 177-184. http://dx.doi.org/10.1145/1013115.1013140
225. CD Wickens and JG Hollands. 2000. *Engineering Psychology and Human Performance*. Prentice Hall.
226. Cary Williams, Xing Dong Yang, Grant Partridge, Joshua Millar-Usiskin, Arkady Major and Pourang Irani. 2011. TZee: Exploiting the Lighting Properties of Multi-touch Tabletops for Tangible 3D Interactions. in *Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11*, Vancouver, BC, Canada, 7-12 May, ACM Press, 1363-1372. http://dx.doi.org/10.1145/1978942.1979143
227. Raphael Wimmer. 2010. FlyEye: Grasp-Sensitive Surfaces Using Optical Fiber. in *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction - TEI '10*, Cambridge, Massachusetts, USA, 24-27 January, 245-248. http://dx.doi.org/10.1145/1709886.1709935
228. Raphael Wimmer. 2011. Grasp sensing for human-computer interaction. in *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, Funchal, Portugal, 22-26 January, 221-228. http://dx.doi.org/10.1145/1935701.1935745
229. Raphael Wimmer and Sebastian Boring. 2009. HandSense: discriminating different ways of grasping and holding a tangible user interface. in *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, Cambridge, UK, 16-18 February, ACM, 359-362. http://dx.doi.org/10.1145/1517664.1517736
230. Jacob O. Wobbrock, Meredith Ringel Morris and Andrew D. Wilson. 2009. User-defined gestures for surface computing. in *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*, Boston, Massachusetts, USA, 4-9 April, ACM Press, 1083-1092. http://dx.doi.org/10.1145/1518701.1518866
231. Katrin Wolf, Anja Naumann, Michael Rohs and Jörg Müller. 2011. A taxonomy of microinteractions: Defining microgestures based on ergonomic and scenario-Dependent

requirements. in *INTERACT 2011*, Lisbon, Portugal, 5-9 September 559-575. http://dx.doi.org/10.1007/978-3-642-23774-4_45

232. Katrin Wolf, Robert Schleicher, Sven Kratz and Michael Rohs. 2013. Tickle : A Surface-independent Interaction Technique for Grasp Interfaces. in *Proc.TEI'13*, Barcelona, Spain, 10-13 February, 185-192. http://dx.doi.org/10.1145/2460625.2460654

233. Fang-Jing Wu, Feng-I Chu and Yu-Chee Tseng. 2011. Cyber-physical handshake. *ACM SIGCOMM Computer Communication Review*, 41 (4). 472-473. http://dx.doi.org/10.1145/2043164.2018527

234. Jiang Xu, Patrick J Gannon, Karen Emmorey, Jason F Smith and Allen R Braun. 2009. Symbolic gestures and spoken language are processed by a common neural system. *Proceedings of the National Academy of Sciences*, 106 (49). 20664-20669. http://dx.doi.org/10.1073/pnas.0909197106

235. Steve Yohanan, Mavis Chan, Jeremy Hopkins, Haibo Sun and Karon Maclean. 2005. Hapticat : Exploration of Affective Touch. in *Proc. ICMI'05*, Trento, Italy, 4-6 October, 222-229. http://dx.doi.org/10.1145/1088463.1088502

236. Xu Zhang, Xiang Chen, Yun Li, Vuokko Lantz, Kongqiao Wang and Jihai Yang. 2011. A framework for hand gesture recognition based on accelerometer and EMG sensors. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 41 (6). 1064-1076. http://dx.doi.org/10.1109/Tsmca.2011.2116004

237. Jamie Zigelbaum, Adam Kumpf, Alejandro Vazquez and Hiroshi Ishii. 2008. Slurp: tangibility spatiality and an eyedropper. in *CHI'08 Extended Abstracts on Human Factors in Computing Systems*, ACM, 2565-2574. http://dx.doi.org/10.1145/1358628.1358713

238. Thoams Guthrie Zimmerman. 1996. Personal area networks: near-field intrabody communication. *IBM systems Journal*, 35 (3.4). 609-617. http://dx.doi.org/10.1147/sj.353.0609