Interaction with Media Facades

The design of interactive systems for large-scale urban screens

Dissertation zur Erlangung des Grades
Doktor der Ingenieurwissenschaften (Dr.-Ing.)
der Naturwissenschaftlich-Technischen Fakultät I
der Universität des Saarlandes

vorgelegt von
Sven Hendrik Gehring, M.Sc.
Saarbrücken, den 13. Dezember 2013

UNIVERSITÄT
DES
SAARLANDES

Dean:
Prof. Dr. Mark Groves

Head of Committee:
Prof. Dr. Andreas Zeller

Reviewers:
Prof. Dr. Antonio Krüger
Prof. Elaine M. Huang, PhD

Committee member:
Dr. Ralf Jung

Day of defense:
February 21, 2014

Acknowledgements

Over the last years I had the pleasure to meet a lot of great people that supported me and my work in various ways. In the following, I want to thank some people in particular and I apologize to the ones I forgot.

First of all, I deeply thank **Antonio Krüger** for the opportunity to work on this thesis and for all the helpful support and guidance. I thank **Elaine M. Huang** for providing me with helpful advice, her support, and for reviewing my thesis. I further want thank **Patrick Olivier** for his valuable support. I also thank **Sebastian Boring** and **Alexander Wiethoff** for great collaborations, fruitful discussions and a lot of fun.

I had the pleasure to work in an inspiring and creative environment. For this, I want to thank my colleagues (in alphabetical order) **Matthias Böhmer**, **Florian Daiber**, **Ralf Jung**, **Gerrit Kahl**, **Christian Lander**, **Markus Löchtefeld**, **Denise Paradowski** and **Johannes Schöning**. I had a great time working with all of you! I am also grateful to the students who shared their distinct and unique abilities to add important bits and pieces to the result of this dissertation. In particular, I thank **Thomas Bauer**, **Magdalena Blöckner**, **Elias Hartz** and **Christian Lander**. Also many thanks to **Margaret De Lap** for proofreading my thesis and to **Martin Zimmer** for sharing his knowledge on 3D modeling with me.

I also want to thank **Sigrid**, **Anke** and **Eva Gehring**, **Rudolf** and **Erna Schneider**, **Gisela Schneider**, and **Gerhard** and **Barbara Hell** for their endless support and for always encouraging and motivating me. Last but not least, I want to thank my wonderful wife **Nicole Gehring** for taking this journey together with me. Thank you for always understanding and supporting me!

Abstract

Media facades are a prominent example of the digital augmentation of urban spaces. They denote the concept of turning the surface of a building into a large-scale urban screen. Due to their enormous size, they require interaction at a distance and they have a high level of visibility. Additionally, they are situated in a highly dynamic urban environment with rapidly changing conditions, which results in settings that are neither comparable, nor reproducible. Altogether, this makes the development of interactive media facade installations a challenging task.

This thesis investigates the design of interactive installations for media facades holistically. A theoretical analysis of the design space for interactive installations for media facades is conducted to derive taxonomies to put media facade installations into context. Along with this, a set of observations and guidelines is provided to derive properties of the interaction from the technical characteristics of an interactive media facade installation. This thesis further provides three novel interaction techniques addressing the form factor and resolution of the facade, without the need for additionally instrumenting the space around the facades. The thesis contributes to the design of interactive media facade installations by providing a generalized media facade toolkit for rapid prototyping and simulating interactive media facade installations, independent of the media facade's size, form factor, technology and underlying hardware.

Zusammenfassung

Die wachsende Zahl an Medienfassenden ist ein eindrucksvolles Beispiel für die digitale Erweiterung des öffentlichen Raums. Medienfassaden beschreiben die Möglichkeit, die Oberfläche eines Gebäudes in ein digitales Display zu wandeln. Ihre Größe erfordert Interaktion aus einer gewissen Distanz und führt zu einer großen Sichtbarkeit der dargestellten Inhalte. Medienfassaden-Installationen sind bedingt durch ihre dynamische Umgebung nur schwerlich vergleich- und reproduzierbar. All dies macht die Entwicklung von Installationen für Medienfassaden zu einer großen Herausforderung.

Diese Arbeit beschäftigt sich mit der Entwicklung interaktiver Installationen für Medienfassaden. Es wird eine theoretische Analyse des Design-Spaces interaktiver Medienfassaden-Installationen durchgeführt und es werden Taxonomien entwickelt, die Medienfassaden-Installationen in Bezug zueinander setzen. In diesem Zusammenhang werden ausgehend von den technischen Charakteristika Eigenschaften der Interaktion erarbeitet. Zur Interaktion mit Medienfassaden werden drei neue Interaktionstechniken vorgestellt, die Form und Auflösung der Fassade berücksichtigen, ohne notwendigerweise die Umgebung der Fassade zu instrumentieren. Die Ergebnisse dieser Arbeit verbessern darüber hinaus die Entwicklung von Installationen für Medienfassaden, indem ein einheitliches Medienfassaden-Toolkit zum Rapid-Prototyping und zur Simulation interaktiver Installationen vorgestellt wird, das unabhängig von Größe und Form der Medienfassade sowie unabhängig von der verwendeten Technologie und der zugrunde liegenden Hardware ist.

Relevant Publications

The work presented in this thesis, including figures and text fragments have in some cases appeared in the following publications. The chapters of this dissertation are partly based on these publications.

Full conference papers:

[66] S. Gehring and A. Krüger. **Façade Map: continuous interaction with media façades using cartographic map projections.** In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp'12, pages 471–480, New York, NY, USA, 2012. ACM. (appears in Section 4.3)

[65] S. Gehring, E. Hartz, M. Löchtefeld, and A. Krüger. **The media façade toolkit: prototyping and simulating interaction with media façades.** In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp'13, pages 763–772, New York, NY, USA, 2013. ACM. (appears in Section 5.2 and Section 5.3)

[194] A. Wiethoff and S. Gehring. **Designing interaction with media façades: a case study.** In *Proceedings of the Designing Interactive Systems Conference*, DIS'12, pages 308–317, New York, NY, USA, 2012. ACM. (appears in Section 4.1 and Section 5.1) **Best Paper Award**

[68] S. Gehring and C. Lander. **GPS Lens: GPS based controlling of pointers on large-scale urban displays using mobile devices.** In *Proceedings of the 2nd ACM International Symposium on Pervasive Displays*, PerDis'13, pages 115–120, New York, NY, USA, 2013. ACM. (appears in Section 4.2)

Short papers and notes:

[**24**] S. Boring, S. Gehring, A. Wiethoff, A. M. Blöckner, J. Schöning, and A. Butz. **Multi-user interaction on media facades through live video on mobile devices.** In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI'11, pages 2721–2724, New York, NY, USA, 2011. ACM. (appears in Section 4.1)

Poster papers:

[**18**] M. Böhmer, S. Gehring, M. Löchtefeld, M. Ostkamp, and G. Bauer. **The mighty un-touchables: creating playful engagement on media façades.** In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI'11, pages 605–610, New York, NY, USA, 2011. ACM. (appears in Section 5.1)

Workshop papers:

[**69**] S. Gehring and A. Wiethoff. **Digital light installations - connecting people through interactive buildings.** In *Workshop on Interactive City Lighting.* ACM International Conference on Human Factors in Computing Systems (CHI-13), Changing Perspectives, April 27 – May 2, Paris, France. ACM, 4 2013. (appears in Section 4.1 and Section 5.1)

[**63**] S. Gehring. **Media facades: turning buildings into large-scale interactive surfaces.** In *Proceedings of the Workshop on "Displays Take New Shape: An Agenda for Future Interactive Surfaces".* ACM International Conference on Human Factors in Computing Systems (CHI'13), Changing Perspectives, April 27 – May 2, Paris, France. ACM, 2013. (appears in Section 1.1)

[**64**] S. Gehring, F. Daiber, and C. Lander. **Towards universal, direct remote interaction with distant public displays**. In *Proceedings of the 3rd Workshop on Infrastructure and Design Challenges of Coupled Display Visual Interfaces* (Capri, Italy), PPD, volume 12, page 302–307, 2012. (appears in Section 4.1)

[**67**] S. Gehring and A. Krüger. **Using media façades to engage social interaction.** In *Adjunct Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp'12, pages 794–795, New York, NY, USA, 2012. ACM. (appears in Section 1.1)

[195] A. Wiethoff and S. Gehring. **Interacting with light.** In R. Wichert, K. Laerhoven, and J. Gelissen, editors, *Constructing Ambient Intelligence*, volume 277 of Communications in Computer and Information Science, pages 63–67. Springer Berlin Heidelberg, 2012. (appears in Section 1.1)

Bachelor's and Diploma theses:

[11] Bauer, T. **Designing Multi-User Interaction with Media Façades**. *Diploma Thesis.* Adviser: Alexander Wiethoff and Sven Gehring, Ludwig-Maximilians-Universität München, Department of Computer Science, 2013. (appears in Section 5.3)

[90] Hartz, E.T. **LASIS - A Simulator for Media Façade Interaction.** *Bachelor's Thesis.* Adviser: Sven Gehring, Saarland University, Saarbrücken, Germany, Department of Computer Science, 2013. (appears in Section 5.2 and Section 5.3)

[111] Lander, C.K. **Distanz-Interaktion mit öffentlichen Displays unter Verwendung eines Mobilgeräts.** *Bachelor's Thesis.* Advisers: Sven Gehring and Markus Löchtefeld, Saarland University, Saarbrücken, Germany, Department of Computer Science, 2013. (appears in Section 4.1)

Contents

1

Introduction

Over the last years, urban environments and public places emerged as prime locations for deploying digital technologies. As can be seen in Figure 1.1, an increasing number of digital systems are rapidly finding their way into these urban spaces [163, 27, 201]. These technologies affect daily life in the public space [121, 180]. For example, they optimize heating systems in buildings, balance the flow of electricity through the power grid, and keep public transportation networks moving [180]. Urban environments are on their way to being turned into dynamic and programmable surfaces [78]. The notion of so-called *Smart Cities* has enjoyed increasing popularity over the last decade. There are enormous efforts by the European Union[1], governments and cities themselves to make cities smart in different domains by utilizing information technologies. The label Smart City is in this sense often used in an ambiguous way. Caragliu defines a city as *smart* when investments in human and social capital and traditional (transport) and modern (ICT) communication infrastructure fuel sustainable economic development and a high quality of life, with a wise management of natural resources, through participatory action and engagement [37]. In general, Smart Cities can be defined in terms of their efforts to create a smart economy, smart governance, smart environments, smart living, smart mobility and smart people by utilizing information technologies [74]. On the one hand, this technology is visible, for example when long-serving artifacts like analog billboards for advertising are replaced by digital displays. Intel estimate in their case study on digital signage that the number of digital public displays will reach 22 millions screens worldwide by the year 2015[2]. On the other hand, a huge variety of sensors and digital

[1] http://www.eu-smartcities.eu/

Figure 1.1: Digital technologies in urban environments.

technologies are deployed into urban areas, not visible at a first glance. For example, surveillance cameras monitor the life in public places and sensor networks monitor the traffic or the air quality. The Smart City of Santander[3] in Spain installed around 3000 wireless sensor nodes, 200 GPRS modules and 2000 joint RFID/QR code labels throughout the city [156].

In addition to the distribution of sensors, the wide distribution of mobile devices brings even more sensors and computing power into the urban environments. Just in the second quarter of 2013, about 435.158.400 mobile devices were sold worldwide[4]. While wireless network connections allow for permanent internet access throughout the urban areas, the deployed technologies are partly connected through various channels. One of the main

[2]https://aimsuite.intel.com/sites/default/files/resources/growing-the-digital-signage-economy-case-study.pdf

[3]http://www.smartsantander.eu/

[4]http://www.gartner.com/newsroom/id/2573415

Figure 1.2: Visual digital impulses in urban environments. Left: Times Square, New York City, USA. Right: Shibuya, Tokyo, Japan.

goals when planning and building an urban space is to achieve that the residents identify with and to relate to the city as their city [161]. Architectural principles state that an effective way of achieving identification is to focus on the communication between people and building [182]. A huge part of the urban space itself is occupied by commercial interests, which usually dictate how the technology is used. The majority of digital displays deployed in public show advertisements dedicated to these commercial interests. To reach as many people as possible, companies want their displays to be present in public spaces, where they are seeking people's attention. With an increasing number of digital public displays, the displays need to compete for maximum exposure [121]. This creates a massive flood of information and digital impulses for the public crowd. Prominent examples are Times Square in New York City, USA and Shibuya, an urban district of Tokyo, Japan (see Figure 1.2). This digital flood of information also affects how people interact with the physical space. The role of public spaces drastically changed from ungoverned places of interaction between people, to controlled spaces, which are artificially designed by city planners to fulfill a particular purpose [62]. Equipping urban spaces with an ever increasing variety of digital technologies creates digital *Mirror Worlds* [70]. But instead of integrating the people into this digital world, they are more and more isolated. The increasing use of digital technologies of various kinds leads to less offline interaction (e.g., face to face conversations) between people [142, 124]. Since the majority of urban displays do not provide any forms of interactivity, the digital impulses created by them to gain the attention of people represent a communication in only one direction. To reinforce human relationships and to integrate humans into the emerging digital cities, we need smart design and a better use of existing technologies to exploit the full potential of the digital technologies deployed in urban areas. With this thesis, we lay the foundation for designing and developing such interactive systems for urban public spaces.

In 2008, the global population reached a historic landmark. For the first time in history, the urban population equaled the rural population [184]. The United Nations (UN) projections indicate that increasing from 3.4 billion people in 2009, by the year 2050, the urban population will expand to 6.5 billion, which will correspond to 70% of the global population [185]. Since with this, the number of potential users of and the number of people that can be reached with digital information technology in urban environments will drastically increase, the careful design of such digital systems will be ever more of importance. In various playful installations, artists and researchers already explored the potential of such digital technologies in an urban context utilizing so-called *media facades* (see Figure 1.1).

1.1 Media Facades & Media Architecture

Besides large scale digital displays, an increasing number of media facades are embedded into the urban landscape (compare Figure 1.3), becoming more and more ubiquitous [191]. So far, there is no clear definition that sufficiently delimits *media facades* from *urban media architecture* and *large scale digital displays*, which are embedded into the urban environment. In a common sense, the term media facade describes the idea of turning the facade of a building into a huge public screen by equipping its outer shell with interactive, light-emitting elements [81, 32, 161]. The display might appear as a second skin of the building. Media facades can be classified based on different characteristics and properties. Among others, these might include their technical composition, as well as the main principles of how media content can be displayed. Besides media facades, the manifold use of light and light-emitting elements in general plays an important role in the architecture of urban environments. In this sense, we have to note the differences between *light architecture, media architecture* and *media facades.* Haeusler [81] distinguishes these terms as follows: Light architecture subsumes the illumination of a building using daylight and artificial light in order to underline parts of the building to create a certain atmosphere. This also holds for media architecture, whereas media architecture also includes all aspects of dynamically displaying media, such as dynamic graphics, dynamic text and spatial movement, but with a strong focus on *dynamic* content. Media facades build on this by including media to transform the building facade into a communicative element, embedded into the architecture of the building. The transition between light architecture, media architecture and media facades can be fluent. Haeusler further argues that a wider definition of architecture often includes the design of the total built environment and the effects on the urban space around media facades [81]. Considering this, he summarizes that media facades embed communication into a facade in the form of digital media, while media architecture describes the cultural, social and economic implications of these facades for the immediate environment [81]. The aforementioned notions of media facades are created from a rather architectural perspective. Since within this thesis we are dealing with media facades as large-scale digital screens from a human-computer interaction perspective, we do not explicitly address the architectural implications of media facades and we therefore do not particularly distinguish light architecture and media architecture. We define the term media facade as follows:

Definition 1. ***Media facades*** *are digital public screens with arbitrary form factors and of arbitrary resolution, which are created by either equipping the outer surface of an architectural building with controllable, uniformly shaped, light emitting elements or by projecting digital content onto it. Media facades are embedded into the architectural structure of the underlying building and transform the building into a communicative element.*

1.1.1 Characteristics of Media Facades

Media facades come with a set of typical characteristics. In the first place, in contrast to situated public displays [137], they are usually very large in size. The size of a media facade can vary from very small media facades of $50m^2$ like the Academy of Fine Arts Saar[5] in Saarbücken, Germany, to medium-sized ones like the ARS Electronica Center[6] in Linz, Austria, covering $5000m^2$, or very large ones, like the Allianz Arena[7] in Munich, Germany, with an area of $25,500m^2$. As a result of their enormous size, media facades can be visible from great distances. This leads to a broad exposure of the content displayed on the facade. In most of the cases, media facades also cover more than one side of a building's facade, in many cases also the roof of the building. This gives them a

[5]http://www.hbksaar.de
[6]http://www.aec.at
[7]http://www.allianz-arena.de

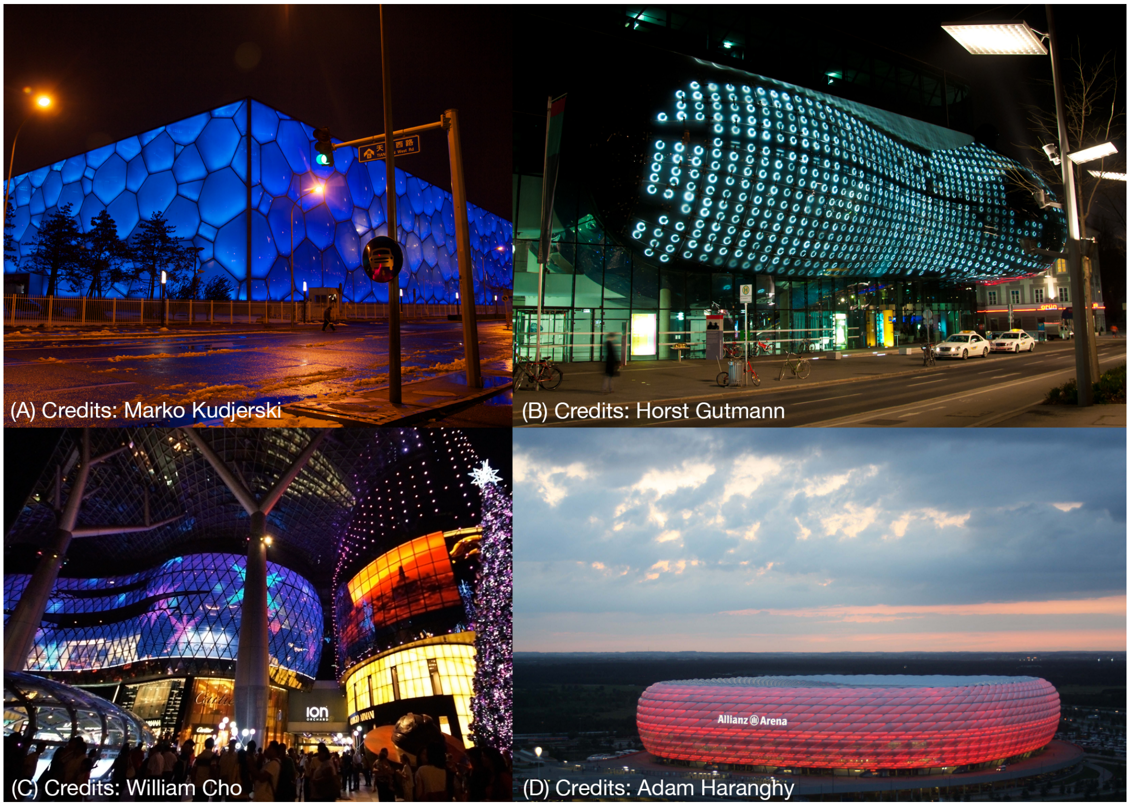


Figure 1.3: Media facades of different sizes and form factors: (A) The National Aquatics Center, Beijing, China, (B) Kunsthaus Graz in Graz, Austria, (C) the ION Orchard building in Singapore and (D) the Allianz Arena in Munich, Germany.

three-dimensional (3D), non-planar form factor. A further very important aspect is the technical specification of a media facade. Media facades are usually individually designed, unique creations. Since it is not common to use standard, off-the-shelf hardware to construct a media facade, media facades are usually very different in their technical setup and therefore, in how to they are accessed and controlled. Furthermore, they also differ in their resolution and hence, in their capability to display particular content.

In their ability to display dynamic, digital media content, media facades might be comparable to situated public displays. Due to their size, unlike situated public displays, media facades require a certain viewing distance to view and comprehend the displayed content. However, situated public displays are a great source of interactivity. People can interact with the display in various ways for different purposes, e.g., looking up information, exchanging content or to entertain themselves. As shown in various works like [145, 137], public displays can also engage people to interact with each other. Due to their characteristics, media facades as well represent an interesting gateway between the personal and the public domain and they offer great potential for interactivity. They provide the possibility to display dynamic digital content and they are visible to a great audience. To establish media facades as such a gateway, media facades need to provide possibilities for interaction and user input.

1.1.2 Interacting with Media Facades

Media facades offer great potential for interactivity. However, this potential has mostly been neglected [177]. Nowadays, media facades are mainly used for displaying pre-produced content, like advertisements or digital animations [94]. In order to exploit the full potential of media facades, we need to develop interactive systems, which allow people to interact with the media facade and also let the media facade engage and mediate the interaction between people. In [177], Struppek investigated the urban potential of public screens. She investigated how the growing infrastructure of digital displays and media facades can be broadened with cultural content. Struppek believes in *interactivity* and *participation* as the keys to bind digital screens to the communal context of the space and thereby, to create local identity and engagement.

Regarding interaction with situated public displays, the most common input modality is direct touch input. Besides the facts that large multitouch screens have become less expensive and direct touch input is an easy to use modality that does not require further hardware or connecting effort from a user, one reason for choosing direct touch input comes with the characteristics of the displays and their deployment context. Situated public displays allowing interactivity are very small compared to media facades, and they are usually installed within the reach of potential users standing in front of the display. Hence, the displayed content is within arm's reach of a user and it is often only visible for people within a — compared to media facades — relatively small distance to the display. For multiple users interacting with the display in parallel, there are spatial constraints for the interaction. At first, since a user occupies space in front of the display while interacting, the number of users is limited by the available space in front of the display. Hence, it is also visible who is interacting with the display and in particular with which parts of the display. As confirmed by [145, 137], this can engage social interaction between users and also between users and spectators in several ways. As a result, conflicts concerning the parallel access of the shared display can be resolved by users directly interacting with each other.

Due to the characteristics of media facades, we cannot transfer these concepts without further effort. As mentioned before, media facades are very large in size and they therefore require a certain viewing distance. Unlike situated public displays, media facades are therefore not suitable for direct touch input, requiring interaction at a distance. Their size, furthermore, leads to a high level of visibility. As a result, the displayed content is visible from great distance and therefore highly exposed to the public. This raises the need for tailored interaction approaches and an adapted design process to address the issues raised by the characteristics of media facades. Along with the technical aspects of media facades, the context in which they are deployed, as well as the exposure of their content have an enormous impact on how users interact with the media facade. Because they are situated in a highly public context, people will adapt their behavior and behave differently when interacting with a media facade than with a situated public display, since they are interacting in front of a large audience and the result of their interaction are highly exposed. Goffman analyzed the behavior of people in public places [75]. He investigated the interaction between people with respect to what happens when two or more people meet. He found that when people meet in public, a system of rules unfolds, which dictates the interaction and which cannot be lead back to the structure and norms of the society. The interaction spans its own realm of interaction. Hence, we should not understand behavior as *behavior* per se, but as a part of an interaction. As described by Goffman, whenever two or more people meet in a public place, a social situation evolves in which people tend to behave in a communicative manner, as if they were interacting with other members of that social situation. This behavior is independent of whether they are interacting or not. This behavior has a significant impact on the way people use interactive systems – like media facades – in a public setting and it therefore needs to be taken into account when designing such systems. In [62], Gehl categorizes activity in public environments into necessary, optional and social activities which accompanied by slightly different behavior. Since the belonging of an activity to one of these categories of activities frames people's behavior, the type of activity needs to addressed when designing interactive systems for public spaces. Furthermore, technology should be designed to enhance and support these types of activities.

1.2 Designing Interactivity for Media Facades

To achieve a usable and enjoyable outcome, designing interaction with media facades demands a structured design process, as well as adapted testing and evaluation methods. The characteristics of media facades, as well as the spaces they are situated in, introduce several design challenges that are critical for successful applications and the enjoyment of their outcome. Besides design process challenges, we have to face further challenges during development and deployment. Opportunities for testing on a media facade during the development cycle are rare. This is due to two reasons: (1) First, so far, most light emitting media facades are not visible and active during daylight, which restricts the available time for testing to only a few hours. (2) In addition, even more importantly, since media facades are situated in a highly public context, everything that is deployed onto the facade is already exposed to a large audience. In this context, testing an early prototype can have an important impact — for better or worse — on how people perceive and experience the final system. A further challenge is the physical properties of the facade and its technical setup. Media facades are mostly unique in their design. They have different form factors and they are custom built with dedicated hardware. For the design and development of digital content, designers usually use dedicated development tools which are specifically tailored to the unique setup of the particular media facade.

1.3 Problem Statement

When designing interactive systems for media facades, the highly dynamic and public environment in which they are situated, and their physical properties, introduce several challenges which need to be addressed in order to create a successful and enjoyable interaction and exploit the full potential of media facades as computing surfaces and large-scale urban displays. In particular, we have to address the following challenges:

- **Categorization:** Each media facade has its unique characteristics, be it the facade itself or the environment it is situated in. There is a need for means to categorize interaction with media facades to allow for comparison and to derive guidelines for future designers to create suitable applications.

- **Design process and tools:** When developing for media facades, we need to develop for a diversity of situations which dynamically change. To represent this in the development process, the design processes need to be tailored to the needs of media facades and new development and prototyping tools need to be developed.

- **Interaction techniques:** Well-developed interaction techniques from related fields like situated public displays cannot be easily transferred to media facades. We need to develop new interaction techniques and adapt existing ones to suit the context of media facades.

To lay the foundation for exploiting the full capabilities of media facades, this thesis addresses the aforementioned problem as follows.

1.4 Research Questions

The research objectives of this thesis are split into three directions. First, a theoretical analysis of interaction with media facade and its framing characteristics, second, interaction techniques for media facades, and third, the design of interactive installations for media facades, which suit the unique properties of these gigantic screens in an urban context. Thereby, this thesis addresses the following research questions:

- **How can media facades be made interactive?**
  To exploit the interactive potential of media facades, we need to investigate approaches for interactivity in different scenarios. Existing interaction techniques need to be adapted and new techniques need to be developed.

- **What are suitable interaction techniques for different contexts?**
  Due to the characteristics of media facades and the urban environment in which they are situated there are special requirements for interaction. We need to identify correlations between interfaces, interactions and emotional qualities when interacting with media facades.

- **What are the key aspects of interaction with media facades?** We need a theoretical analysis of existing media facade installation in order to identify the key properties framing the interaction. We need to derive a categorization based on these characteristics to provide means for making interactive media facade installations comparable with each other to inform the design and development process.

- **How to design interaction for media facades?**
  To achieve a usable and enjoyable outcome, we need a structured design process and generalized design tools. This design process needs to address the dynamic, public context of media facades. Furthermore, we need to focus on facilities of prototyping to overcome the existing shortcomings.

1.5 Methods & Approach

The research described in this thesis follows an exploratory, user-centered approach [51, 130, 155, 166], both on an experimental and theoretical level. We build high- and low-fidelity prototypes involving users to ensure their needs are met and create a usable and enjoyable outcome. The needs of the users are addressed in the particular stage of the design process by incorporating repetitive feedback loops. To obtain a better understanding of how potential users perceive initial design ideas, we conducted a series of interviews *in the wild*. This provides valuable new insights and perspectives from which the further stages of the development process can benefit.

Besides collecting feedback on the usability of the prototypes, we focus in particular on making use of user experience (UX) [8] in combination with human-computer interaction (HCI) [51] evaluation methods to investigate the particular prototypes more holistically.

1.6 Contributions

The goal of this thesis is to take the first step towards turning media facades into fully fledged interactive surfaces which enrich the urban landscape and offer new perspectives for urban computing. To achieve this, we aim at contributing in three aspects:

1. **Theoretical contribution:** Up to now, there is little theoretical understanding of which parameters define and shape interaction with media facades. Furthermore, there is neither common ground on the suitability of different interaction techniques in particular settings, nor any categorization that puts the components of media facade installations into context. We contribute to this by conducting a theoretical analysis of the design space for interactive installations for media facades. We investigate (1) the extent to which an interactive installation allows passersby to participate, (2) the resolution of the media facade itself with respect to the resolu-

tion of user input, as well as (3) the technical characteristics of the overall setting. We provide taxonomies for each of these three aspects. The provided taxonomies put media facade installations into context and allow for comparison. Furthermore, we introduce metrics accompanying the taxonomies: first, a resolution factor that relates input and facade resolution to indicate the suitability of an interaction technique to allow for a particular granularity of interaction for a media facade. Second, a set of observations and guidelines to derive properties of the interaction from the technical characteristics of an interactive media facade installation, such as for example the size of interaction space or appropriate interaction distances (see Chapter 3).

2. **Technical contribution:** In order to interact with media facades as interactive surfaces, we need interaction techniques allowing for fine grained interaction with the media facade's content with respect to its resolution and shape. As a technical contribution, we introduce three novel techniques to interact with a media facade at a distance. All three techniques allow multiple users to interact simultaneously. They utilize regular mobile devices and do not require an instrumentation of the potential interactive space around the media facade with additional technology. This makes them applicable in a wide range of settings. They further address the form factor and the resolution of the facade as characteristic properties in particular. First, with iRiS, we provide a technique to interact with media facades of arbitrary resolution by allowing for interaction through live video on mobile devices (compare Section 4.1). This approach especially contributes to interacting with low resolution facades, since visual feedback can be augmented in the live video on the mobile device, which does not waste screen real estate of the facade. Second we provide the GPS Lens, a sensor-based interaction technique to control a pointer on a media facade by pointing at the facade with a mobile device (see Section 4.2). Third, we provide a technique to interact with media facades of arbitrary form factors using cartographic map projections on a mobile device, where users interact with a cartographic map of the media facade's surface (see Section 4.3).

3. **Design contribution:** We contribute to the design of interactive media facade installations by providing a generalized media facade toolkit for rapid prototyping and simulating interactive media facade installations. The toolkit can simulate large-scale media facades, independent of the media facade's size, form factor, technology

and underlying hardware. Furthermore, it allows to run arbitrary applications on arbitrary media facades, which lowers the strong bond between application and media facade that exists with current development toolkits. With this, we allow for the transferability of applications from one media facade to another, which was so far one of the major disadvantages of current development tools. Applications can be mapped onto media facades as texture images by integrating cartographic map projections into the toolkit to determine texture coordinates for arbitrarily shaped surfaces. With the media facade toolkit, we further make it possible to simulate different hardware, media facade resolutions and different color models. We address the diversity of media facade settings and incorporate the media facade's architectural surroundings into the simulation (see Chapter 5).

1.7 Thesis Outline

The remainder of this thesis is structured as follows: In Chapter 2, we give a detailed overview on related literature that is relevant to this work. The overview on related work is followed by a theoretical analysis of interactive media facade installations in Chapter 3. We derive a taxonomy to categorize existing and future installations based on characteristic properties. In Chapter 4, we introduce new techniques for interacting with media facades that fill the gaps among existing approaches according to the categorization provided in Chapter 3. Chapter 5 addresses the process of designing interaction for media facades. We review our experiences from developing different prototypes as the foundation for the design of a simulation tool, which allows prototyping and the simulation of interactive systems for media facades under dynamic conditions within a 3D environment. The thesis concludes by summarizing the main contributions of this work and identifying future work for which the results of this thesis can lay the foundation.

The major part of the work that is presented within this thesis was carried out in collaboration with researchers, students and professionals from various institutions. For this reason, the scientific plural "we" is used throughout the thesis.

2

Background & Related Work

The work presented within this thesis builds upon the areas of *shared encounters in public spaces*, *interaction with public screens* as well as *designing and prototyping interaction for public spaces*.

In this chapter, we illuminate the topic of media facades as interactive shared encounters in urban spaces from different perspectives. Therefore, we split this chapter into three main parts. In the first part of this chapter, we review media facades as digital installations in urban spaces. Within this scope, we address their role as shared encounters, offering great potential for interactivity. We investigate the role of the environment media facades are deployed in, as well as the configuration of the spatial setting in front of an interactive media facade. Furthermore, we compare the key features of media facade installations to the related field of situated public displays, outlining their differences and their commonalities. In the second part of this chapter, we address interaction with urban screens in general, ranging from situated public displays to large-scale media facades. We provide an overview on existing models describing behavior around and interaction with such digital public screens, as well as examining recurring patterns concerning behavior and design. This is followed by an overview on interactive media facade installations and relevant techniques for interacting with digital screens at a distance with respect to their applicability for media facades in a public context. The third part of this chapter is dedicated to designing, prototyping and evaluating interactive systems for public settings. We outline the current design practices and challenges, as well as their implications for this thesis. Within this scope, we further address evaluation methods for analyzing interaction in public spaces, addressing their unique characteristics.

2.1 Media Facades & Public Spaces

Media facades not only represent a novel type of interface, but also a great source for interactivity. Due to their physical and digital properties, as well as the public setting they are usually situated in, we have to face novel challenges when designing and developing digital content for media facades. Hence, we split this section into two parts. The first part describes work about media facade and media architecture installations with respect to the general setting and their meaning for the design of interactive systems. The second part gives an overview on existing approaches to realize interactive media facade installations.

The recently emerging field of *Urban Computing* is addressing the increasing availability of digital technologies in urban spaces, as well as their use. The umbrella of Urban Computing covers an interdisciplinary field, bringing together art, architecture, urban planning, geography, social sciences and computer science [143]. The combination of architecture and media or display elements has been extensively investigated throughout the last decade. Haeusler, Bullivant and Schoch each described how the facade of a building can be turned into a huge public screen by equipping its outer shell with interactive, light-emitting elements [81, 32, 161]. As mentioned in the previous section, Haeusler provides a first approach to categorize existing media facade installations with respect to their technical realization and he gave a first formal definition of media facades from an architectural perspective [81].

Within the scope of Smart Cities, media facade are considered as important elements to connect the traditional city with its landscape and architecture to the digital elements driving the city. Frenchman and Rojas describe the role of media facades and media architecture in making places within a Smart City that respond to people, change to accommodate multiple activities or provide stories, information and services [59]. Within this scope, they consider the traditional boundaries between the physical and digital parts of a city to melt into a layered cityscape of fine-grained places, visual images, and multiple shifting activities [59]. Kroner sees digital enhanced buildings as a forerunner for the architecture of future cities [108]. He states that the resulting architecture will be comprised of responsive environmental systems and will mark the emergence of the *Knowledge Building*, a fully developed intelligent architectural form with media facades as the connection to the world around the building [108]. Stalder also considers media facades as an architectural enhancement and a prime location for displaying pervasive advertisements in future cities [175].

In [161], Odilo Schoch describes the possibilities and the potential of using buildings or parts of their outer surface as multidimensional graphical output devices. He outlines how such novel types of interfaces differ from those of personal computers and that their ability to handle a large number of users challenges the displayed content. He further argues that architects are challenged by aesthetics, technological and social issues when designing for media surfaces. Schoch also addresses interaction with media facades, stating that novel types of interaction can become mandatory, as a display in urban scale is different from a personal computer [161]. This requires further research to investigate novel interaction techniques and to adapt existing ones, since the interaction is carried out in a new, highly public setting. This further raises the need for suitable development tools and processes to address the requirements of this novel domain. Mignonneau and Sommerer picked this up and further explored the social potential of media facades as digital screens that can be seen or even designed by multiple persons simultaneously [127]. They describe media facades as allowing building facades to become membranes for the display of interactive digital content, which they argue raises new challenges for the design [127]. They believe media facades to be social mediators of interaction and they see interactivity as an important factor for that.

Struppek investigated the urbane potential of public screens in general with respect to the urban public space, to which she refers as *a key element in the development of European urbanism* [177]. She investigated how the growing infrastructure of digital displays and media facades, which is currently dominated by commercial interests, can be broadened with cultural content. She argues that architectural aspects of urban spaces play an important role in providing a stage for interactions, where the architecture itself serves as a mediator. Media architecture and media facades as new elements in the urban space come with new possibilities for interactivity and participation. Struppek argues that interactivity and participation are the keys to binding these digital screens to the communal context of the space and with this, they create local identity and engagement. In [109], Kuikkaniemi et al. investigate how the vastly growing number of interactive urban screens in general will change urban life in the future. They argue that with the massive deployment of display technologies of various kinds, nearly any surface in urban environments might become a digital display. This will lead to walls and building facades which creatively motivate interaction, both individual and group interaction. According to Kuikkaniemi et al., digital technologies and media content massively brought to urban environments will reestablish public places as the center of social life, where it actually originated.

They way how media facades are used, how they are perceived and possibilities they offer are heavily affected by the properties of the environment, in which they are situated. First, the public nature of the deployment environments plays an important role. The media facade and hence its content are highly visible and exposed to a large audience. Second, since interaction wit a media facade usually takes place in the area in front of the building, the spatial configuration of the deployment environment has a strong influence.

2.1.1 The Spatial Setting

Fischer and Hornecker analyzed the spatial aspects in the design of shared encounters for interactive media facades [54]. They reflected on various urban technology interventions by analyzing their spatial configuration in relation the structuring of interaction. They introduce the term *Urban HCI*, which focuses on urban settings where context is not only a location point but also activity. Urban HCI is utilized to emphasize situations composed of the built environment, the interfaces and the social context. Fischer and Hornecker identified the following seven spaces, which are depicted in Figure 2.1. The *Display Spaces* are all spaces from which a media facade is visible. Considering the enormous size of media facades, compared to common situated public displays, the Display Space can be huge, ranging from a plaza in front of the facade to whole parts of a city. As *Interaction Space*, they describe the space from which a form of communication with the media facade installation is carried out. Size and placement of the Interaction Space depend on the applied interaction techniques and whether it is stationary or mobile, allowing users to walk around and to interact from various places. The *Potential Interaction Spaces* are the spaces from which a person could potentially interact with the media facade. The size of the Potential Interaction Space also depends on the properties and the range of the applied interaction technique. *Gap Spaces* can create distances and gaps, either between people or between people and facade. The gaps in this case can be introduced by natural

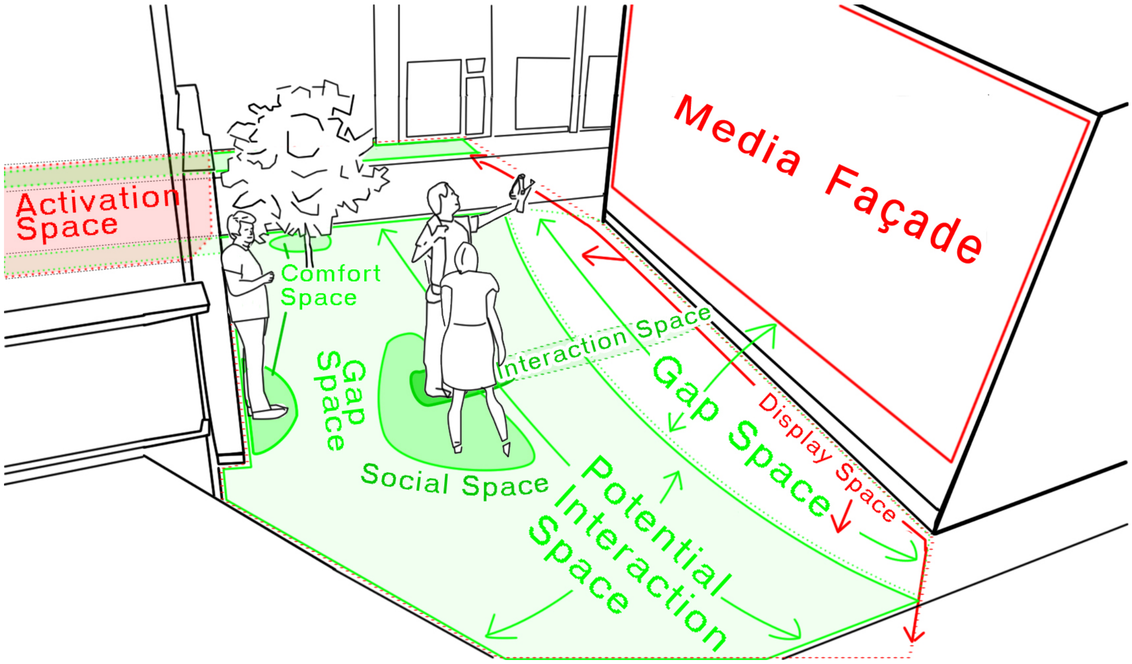


Figure 2.1: Different space types for media facade installations according to Fischer and Hornecker [54].

or artificial obstacles (e.g., trees, streets or benches) and also by the required minimal viewing distance to perceive a sufficient part ot the facade's content. Spaces where people gather and where they can have shared encounters without necessarily interacting with the facade are denoted as *Social Interaction Spaces*. In *Comfort Spaces*, people can find physical or psychological ease. Comfort Spaces can occur around objects like trees or walls of surrounding houses, where people can opt out of the social setting and blend into the environment in order to observe. Finally, *Activation Spaces* are the areas from which the media facade is visible, possibly causing curiosity and triggering passersby to approach and participate. For the remainder of this thesis, we use the terminology introduced by the Urban HCI model [54] when describing media facade installations.

Understanding the spatial configuration of the spaces around a media facade is an important factor for the design and development of successful interactive installations for media facades. When creating such systems, the spatial configuration of the media facade's environment further influences the appropriateness of certain content and it further restricts the applicability of different interaction techniques.

Figure 2.2: Public displays of various sizes.

2.1.2 Media Facades versus Public Displays

Although media facades and situated public displays seem to be closely related as they both share the purpose of being public screens, communicating digital information in a public environment, there are elementary differences between both. In order to clarify the varying approaches for designing content and applications for both, we highlight the key differences:

Integration into architecture

The integration into architecture is not only important from an aesthetic point of view. It makes a crucial difference between public displays and media facades. While Schoch distinguishes between dedicated *media facades* and *buildings designed with media technology as a main element* [161], Tscherteu combines the both [182]. However, Tscherteu and Schoch both clearly distinguish light and screen elements integrated into the architecture of a building from so-called *urban screens* as add-on displays, such as situated public displays. Tscherteu draws the boundary where it is no longer the screen on a building that communicates with its surroundings, but rather the building as a whole [182]. The integration of screen elements into existing architectural structures often also results in uncommon, irregular form factors in contrast to the usually rectangular-shaped situated displays with a 2D form factor.

The spatial setting

As shown by Fischer and Hornecker [54], the spatial aspects of media facades have a strong impact on the installations. Describing discrete spaces around an interactive media facade, they clearly distinguish media facades from situated public displays. Müller et al. [131], analyzing requirements and the design space for situated public displays, as well as Diniz et al. [50], investigating territoriality and behavior within the scope of interactive media facades, further indicate the spatial setting around the screens as an important difference between situated public displays and media facades. While for situated public displays people usually are within an arms length from the display when interacting, for media facades, the interaction is performed within a wide, distributed space in front of the facade. Interactions in such a case are shown to have different characteristics with a strong focus on territoriality and personal space [54, 50].

The scale

A further fundamental difference between media facades and situated public displays is the physical dimension of the screens. While the size of public displays ranges from TV- to billboard-sized screens (as in Figure 2.2), media facades usually achieve architectural scale. E.g., the previously mentioned media facade of the Allianz Arena in Munich, Germany, has a surface of $25.500m^2$. Such enormous sizes affect the visibility of the content and the way people interact. It requires a certain minimal viewing distance to perceive the displayed content, which introduces gap spaces [54]. However, as can be seen in the example of the Allianz Arena, media facades can be so big that it is impossible to see the whole facade at once. Furthermore, people interact differently. While for situated public displays users can see each other since they are close to the display, for media facades, users can be distributed through large areas, not aware of other users. This circumstance influences the social protocols of how people behave [75].

Diversity of situations

Just like media facades, public displays are usually situated in a dynamic public environment. While the space around the display is rather limited, the number of people around the display, as well as the general conditions of the environment can vary. However, for media facades the diversity of situation is way more distinctive. Due to their size and the fact that they are located in outdoor urban spaces drastically increases the number of variables in the setting. Changes in weather and lighting conditions, population of the surrounding space, traffic are only a small number of factors influencing the whole setting. They have an impact on how people approach, perceive and interact with a media facade. For media facades, it is barely possible to reproduce a complete setup or a particular situation. This diversity of situations needs to be addressed when designing content and interaction for media facades, as when providing the respective development tools.

2.1.3 Shared Encounters

Creating engagement in urban spaces is of great importance when dealing with media facades as *shared encounters*. Schieck at al. refer to the concept of shared encounters as an *an ephemeral form of communication and interaction augmented by technology* [71]. Willis et al. further define it as the interaction between two people or within a group where a sense of performative co-presence is experienced by mutual recognition of spatial or social proximity [196]. Fischer and Hornecker consider shared encounters as short intermezzos in our habitually predetermined everyday urbane life, or dérive creating works that are integrated in the built environment, without demands of permanence in time [54]. According to them, for urban environments, the concept of shared encounters brings together architecture, urbanism, social sciences, anthropology and computer science. When creating interactive installations in urban spaces, an important aspect is to create a binding between the environment and its inhabitants. To address this, besides technical aspects, understanding and creating engagement is an important need.

Dalsgaard et al. analyze three of their previous media facade installations in terms of their support of engagement as an evolving process. They specifically focus on how engagement unfolds as a dynamic process that may be understood in terms of evolving relations between cultural, physical, content-related, and social elements of interactive environments [44]. They consider engagement as a perspective on interaction, rather than a clearly defined entity. Through reviewing literature on engagement and interaction design, they extracted various elements of experiencing engagement, covering physical, cultural, content and social elements. They conceptualize engagement as *evolving relationships between physical, social, cultural and content-oriented elements* and they outline that understanding the dynamics of engagement with any given interactive installation entails understanding how these elements are continuously re-shaped and formed into new constellations [44].

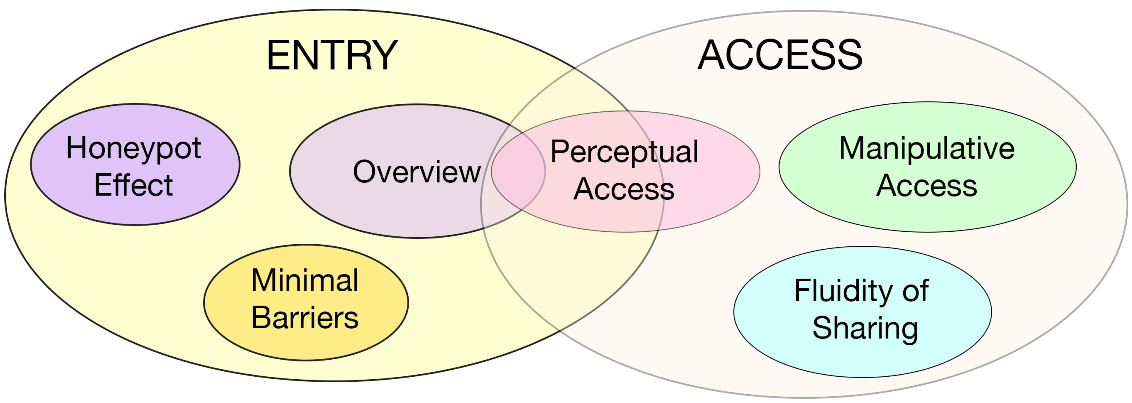


Figure 2.3: The relationship between elements of entry and access points for shareable interfaces according to Hornecker et al. [95].

Hornecker et al. introduce the design principle of shareability, which they refer to as *a design principle that refers to how a system, interface, or device engages a group of collocated, co-present users in shared interactions around the same content or object* [95]. They break participation down into the notions of entry points and access points, where entry points invite and entice people into engagement, providing an advance overview, minimal barriers, and a honeypot effect that draws observers into the activity; and access points that enable users to join a group's activity, allowing perceptual and manipulative access and fluidity of sharing [95] (see Figure 2.3). One of their main findings is that the design concept of shareability is useful in the analysis of existing systems and in framing experimental questions about the properties of particular shareable systems. They believe that it can help designers to think about trade-offs when designing interactive systems for multiple people.

Brignull and Rogers investigated engagement and shared encounters, specifically how to entice people to interact with situated public displays in a public setting [28]. Enticing people to engage is a widespread issue with interactive installations in public environments. Brignull and Rogers found embarrassment as an important factor. In accordance to Goffman [75], they found that people tend to feel watched and embarrassed when interacting in front of a group of other people. For designing interactive systems for public spaces, they propose to aim for strong physical and social affordances and to design ways of encouraging people to cross the threshold from peripheral awareness to focal awareness, to participation and back again, without becoming self-conscious [28].

Investigating a similar setting, Schieck et al. explored shared encounters mediated by technologies in urban environments [72]. They deployed an interactive public display in an urban environment and observed the behavior of passersby with a focus on aspects influencing interactions of people when technology is introduced into the setting. The aim of their work was to create novel experiences that trigger shared encounters among people. Their observations revealed that introducing such a digital system made people aware of the existence of others in the same area. They concluded that in reference to Goffman [75], this awareness can influence the behavior of the people and that this can motivate them to change the way they communicate and engage with others, creating shared encounters.

Harrison and Dourish analyzed different notions of *space* and *place* used to facilitate and structure interaction in public settings [89]. They argue that the common focus on spatial models is misplaced and that we have to critically distinct between space and place. While space as a fundamental concept in architecture and urban design denotes the spatial arrangement of a particular real-world setting, of which we have a mutual understanding, the concept of place derives from a tension between connectedness and distinction. Our sense of place is a cultural or communally-held understanding of the appropriateness of styles of behavior and interaction, which *may* be organized around spatial features [89]. They argue that in our daily experiences and interactions, it is the sense of place and not the spatial arrangement of the space which frames our behavior. Hence, the notion of place turns spaces into shared encounters. In short, the principle of Harrison and Dourish can be summarized as: *Space is the opportunity; place is the understood reality* [89]. Dourish revisited the notion of space and place a decade after the initial publication [52] showing that the concepts have proven useful across a wide range of different domains. However, after significant developments in spatial technologies, in particular the adoption and widespread use of wireless and mobile systems, he shifts the focus back toward space and outlines that it is important to see both as *critical aspects and products of the circumstances of the interaction.*

The described social factors for spatial settings in public spaces, such as interaction with media facades, are of great importance for the design of enjoyable and successful interactive systems. They need to be addressed during the design process, to tailor the developed system to the needs of the actual users.

2.2 Interacting with Urban Screens

As mentioned before, interactivity and participation are keys to binding digital screens to the communal context of the space and with this, to create local identity and engagement. However, urban screens providing interactivity are still rare. With the umbrella term urban screens, we summarize digital displays deployed in urban public spaces. This includes situated public displays, digital signage, video walls, as well as media facades. These displays are mostly used for displaying advertisements, general information, or static content. For situated public displays, if they are interactive, the most common interaction technique is manipulation by direct touch input, which is only applicable if the user can get close enough to the display to actually touch it. Due to their enormous size, media facades require interaction at a distance, since interaction by direct touch is not applicable. Existing approaches allowing for interaction with media facades either use custom-made interaction techniques and input devices, or they follow a playful approach. In both cases, the interaction is highly tailored to the particular setting and therefore not directly portable to other media facades. To be able to establish media facades as the future display surfaces in urban spaces, we need find suitable *general* interaction techniques that are easily portable. Therefore, we need to both develop novel interaction techniques suitable to address the criteria imposed by media facades, as well as to adapt existing techniques to interact with situated displays at a distance.

In the following, we review relevant work on interactive media facade installations based on the applied input modality, as well as relevant approaches for interacting with situated displays at-a-distance with respect to their applicability to media facades.

2.2.1 Models, Patterns & Guidelines

In current research we can find several *models* describing interacting with various public displays, *patterns* for the behavior of people around such displays and for the design of the respective systems, as well as guidelines for the design of interactive installations for public settings. While models are formalized holistic descriptions of a phenomenon, a setting, a sequence of actions or the general behavior of people approaching a display and interacting with it, patterns describe recurring structures in both particular behavior of users and design structures. They help to understand the behavior and interactions of people in the respective settings and they provide useful guidelines for developers, developing interactive settings for public environments. Although not being the main scope of this thesis, they inform the design and development of techniques to interact with public displays and media facades. Hence, we will summarize the relevant models, patterns and design guidelines in the following.

Models

With the aforementioned *Urban HCI* model, Fischer and Hornecker introduce the first model to describe the spatial setting around interactive media facades [54]. They analyze the spatial configuration of interactive media facade installations in relation to the structuring of interaction. The Urban HCI model is utilized to emphasize situations composed of the built environment, the interfaces and the social context. It provides different spaces to describe the spatial configuration of an interactive media facade setting (compare Section 2.1.1). With this, the model can be applied to predict spaces from where people can potentially interact with the media facade, in which spaces social interactions are likely to happen and from where bystanders most likely observe the interactions. The model not only allows to describe and analyze existing situations, but it also provides a terminology that can be applied when planning new interactive installations. Disciplined thinking about different spaces and display types can inform design strategies for creating new installations and may provide guidance when moving installations to new locations [54].

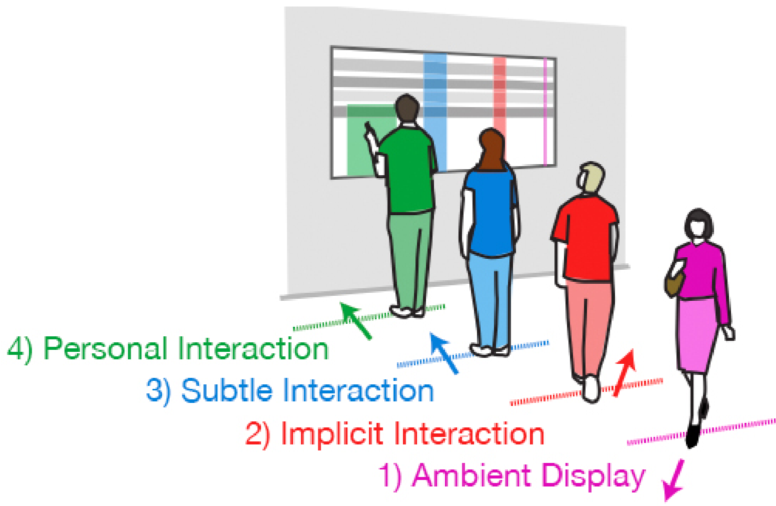


Figure 2.4: The four interaction phases according to Vogel and Balkrishnan [187]. They are facilitating transitions from implicit to explicit and public to personal interaction.

The Urban HCI model is the first formal model describing the spatial setting in front of interactive media facade installations and providing a terminology to describe the setting. Regarding public displays, there exist further models describing the spatial settings in front of a situated public display, as well as how people perceive and recognize such displays within their environment. These models have a different scope and they are dedicated to much smaller and more closed settings, for which they are not immediately transferable to media architecture and media facades. However, for the sake of completeness, we will briefly mention the most prominent models. They have different emphases in terms of roles, territory and behavior.

One of the first models describing how people interact with a shared public display was introduced by Vogel and Balakrishnan [187]. The model comprises an interaction framework and it describes the transition from distant, implicit, to explicit personal interaction with four interaction phases that can be seen in Figure 2.4:

1. The *Ambient Display Phase*,
2. the *Implicit Interaction Phase*,
3. the *Subtle Interaction Phase* and
4. the *Personal Interaction Phase*.

A user can seamlessly switch between the different phases. Furthermore, the model also describes how the granularity of the displayed content should be adapted to the particular interaction zone of the user, ranging from ambient content to explicit personal content. The Ambient Display phase is the neutral state of the model where the display serves as an ambient information displays. In the Implicit Interaction Phase, the display shows peripheral notifications when a user passes by, in order to raise his awareness for the display. When the user then approaches the display, he enters the Subtle Interaction Phase, where the display is showing a more detailed description of the notifications and the information available on this display. By starting to directly interacting with the display, the user enters the Personal Interaction Phase, in which the display shows more details, including personal information. In contrast to the models described in the following, the model by Vogel and Balakrishnan was not derived from observations how people approach and use public displays in the wild. It was developed and presented as a part an interaction framework, providing means to raise the awareness of a user for a public display and to engage him to approach the display and interact with it.

Kaviani et al. introduce a *Role Centered Model* to categorize the audience of a public display into roles describing the extent to which they interact with the display [101]. They define a number of design strategies focusing on the concepts of interactive public displays. As depicted in Figure 2.5, they categorize the audience of a public display into *actors*, *spectators* and *bystanders*. Actors actively engage with the display. They directly interact with the display and they actively manipulate content. Spectators are mentally engaged with the display and its environment. While spectators do not actively interact with the display and its content, they watch and observe actors while they interact. Bystanders are people located in the area around the display, having no interest in the display, the interaction or the displayed content. When designing interactive applications for public displays, this model can help to both analyze the behavior of the users and to design interaction and content addressing the particular user roles.

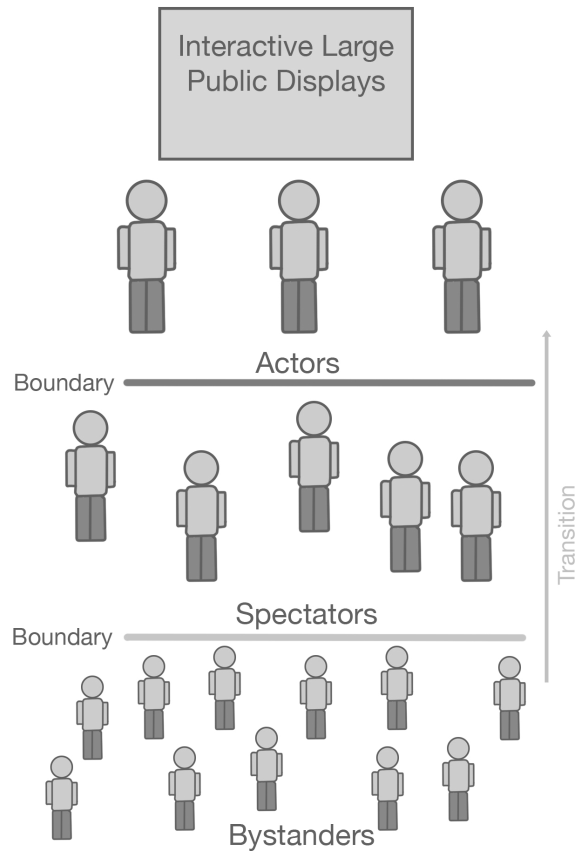


Figure 2.5: The role-centered model for public displays according to Kavani et al. [101]. The audience in front of the display is categorized into actors, spectators and bystanders.

With the *Audience Funnel*, Müller et al. provide a comprehensive analysis of the design space explaining mental models and interaction modalities from public displays [131, 126]. On this basis, Michelis and Müller derived the Audience Funnel model [126]. Observations of people's behavior when interacting with "Magical Mirrors" revealed recurring behavioral patterns, such as glancing at the display while passing it or coming closer to interact. From these observations they deduced a model of interaction — the *Audience Funnel* — with gesture-based public display systems. As depicted in Figure 2.6, the model comprises the phases *Passing By*, *Viewing and Reacting*, *Subtle Interaction*, *Direct Interaction*, *Multiple Interaction* and *Follow-up Actions*. When interacting with the display, people pass through the particular phases, overcoming respective thresholds in each phase. Since this model described how people approach interactive public displays, it can inform the design of interactive applications for public displays in terms in terms of designing the content in a way such that the passersby are engaged to stop and interact.

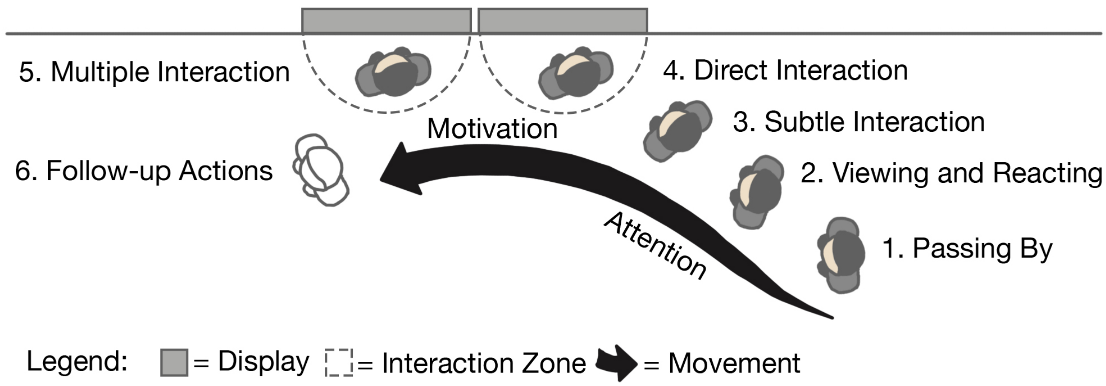


Figure 2.6: The Audience Funnel [131, 126] describing the phases of interacting with a public display.

With the *Activity Spaces* model, Brignull and Rogers put more focus on the social setting of the space around the display rather than the space itself [28]. They consider the often observed resistance by the public to participate and interact with a public screen in front of other people as a major problem. As the main reason, they consider the prominence of the affective aspect of the user experience, where feelings of social embarrassment often act as a barrier [28]. They present a model of public interaction flow, from which designers can derive design recommendations for encouraging public participation. The model comprises of three distinct activity spaces:

1. *Peripheral Awareness Activities* are typically activities like eating, drinking or socializing in the space around the display. People in such Peripheral Activity Spaces are peripherally aware of the display's presence, but they do not have further knowledge about it.

2. In the *Focal Awareness Activity* space, people pay more attention to the display and how it is used. They are usually engaging in socializing activities with respect to the display. These include gesturing to and watching the display being used or also talking about it to others.

3. In the *Direct Interaction Activity* space, people are directly interacting with the display.

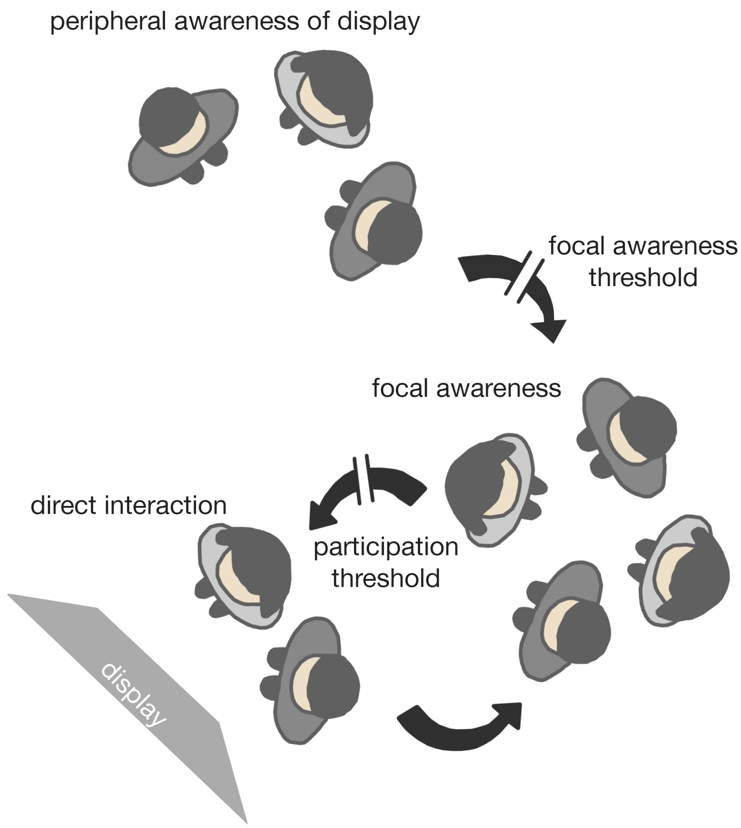


Figure 2.7: Brignull and Roger's model of interaction flow across thresholds in front of a public display [28]. The interaction flow transits from peripheral display awareness over focal awareness to direct interaction with the display.

Brignull and Rogers further point out, how their model informs the design of interactive applications for public displays. They derive guidelines for how displays should be positioned in public spaces and how content should be designed to engage passersby to overcome the thresholds and interact with the display.

With the *Passive Engagement, Active Engagement, and Discovery* model — in short the *PACD* model — Memarovic et al. divide the space in front of an interactive public display into different zones of engagement [125] (see Figure 2.8). The model describes multi-user coordination and engagement around the display and it was derived form observing people interacting with the FunSquare application that displayed locally scoped "fun facts" [123]. It consists of the (1) *passive engagement zone* and the (2) *active engagement zone* and describes how user activities within the active engagement zone influence activities and behavior in the passive engagement zone. Within the passive engagement zone, people can have *glimpse interactions* with the display, where they observe other users, or they can have short interactions with the content, such as "read and go". They can further have *immersive interactions*, where they interrupt their current activity while walking by and pay attention to the display. In the active engagement zone, people actively engage and interact with the content displayed on the public display in a more focused way. These focused interactions contain *active reading* of the content, *read'n'interact* (i.e., reading content items shown on a display and interacting with them) or *direct interaction* where the users are interacting with a display prior to reading the displayed content. The PACD model complements model of Brignull and Rogers [28] and the Audience Funnel of Michelis and Müller [131, 126] by describing turn taking in front of a display and how active engagement with a display sparks activities and interest of passersby.

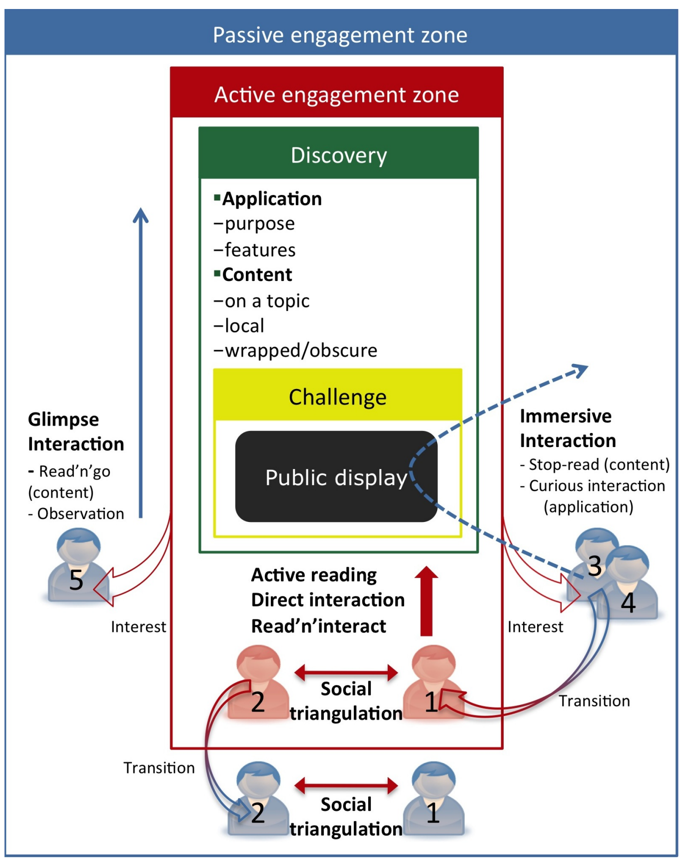


Figure 2.8: The *Passive Engagement, Active Engagement and Discovery* (PACD) model conceptualizing engagement zones and activities around public displays [125].

The aforementioned models mostly concern interactivity with situated public displays. In such scenarios, the space in front of the display, and hence the number of people involved are rather limited. For media facades, it remains to transfer existing models to the scale and complexity of the interactive spaces around a facade. Fischer and Hornecker [54] and Vande Moere et al. [129] tried for the first time to model the space in front of and the behavior of people interacting with, a media facade. Although the scale and complexity of the setting required adaptation of existing models and it introduced new spaces and behavior, it turned out that both worlds — media facades and situated public displays — have their similarities and share recurring patterns. To follow we will identify shared patterns collected from existing work.

Patterns

From interactive installations for media facades and interactive public displays, we can identify a number of recurring behavioral and design *patterns* that are present for both scenarios. In contrast to models which holistically describe processes or settings, patterns are discernible regularities that can be observed. These patterns can inform a better understanding of interactive installations in public, and they can also help developers to design interactive installation aiming for a certain behavior. From literature, we extracted the following recurring patterns:

Entry Points and Access Points is a pattern introduced by Hornecker et al. [95] dealing with how people engage with interactive installations. It motivates the design of entry points to create engagement and make people aware of an installation, as well as access points, which are characteristic elements allowing people to interact with an installation either on their own or joining a group activity.

The **Honey Pot** effect described by Brignull and Rogers [28] describes the phenomenon of a progressively increasing number of people in front of a public display, creating a *social buzz* around the installation. By standing in the area in front of the display, people seem to send an implicit signal to others that they are open to conversation and interaction.

Related to the Honey Pot effect, Dalsgaard et al. introduce the **Initiating, Transforming, Resuming** pattern [44], which comprises a set of five sub-patterns as described by Brynskov et al. [30]: (1) With *Watch-and-join*, they refer to interactions that are initiated by people first watching an installation and then joining people currently interacting with it. (2) Similar to Watch-and-join, *Watch-and-take-over* denotes situations where people initially watch the installation and people interacting with it. But here, instead of joining current users, people wait for the current users to finish and then they take over. (3) With *Walk-up-and-use*, Dalsgaard et al. denote situations where a person approaches an installation an immediately starts interacting with it. (4) *Interact-and-run* describes a similar pattern, where in contrast to Walk-up-and-use, a person approaches the installation and only briefly interacts before leaving. (5) Finally, returning the installation or handing it over after interacting with it is denoted with *return*.

Jacucci et al. describe the **Social Learning** pattern [97]. It maps behaviors people employ to understand an installation and to learn how to use it. *Individual exploration* describes one person testing out and exploring the installation without paying attention to other users. For *cooperative exploration* people get together in pairs or groups of people to explore the installation together. The behavior of watching other users interacting with the installation and then trying to imitate their strategies or trying their own ones after it is denoted as *passive observation then attempt*. Finally, *imitation* denotes when people directly approach the installation and start to imitate the interactions of present users while they are there.

With the **Active Spectatorship** pattern, Peltonen et al. follow a different perspective on spectators of an interactive installation in a public space [146]. Their concept aims at extending the feeling of participation in a large-scale event transforming the rather passive role of a bystander into what they call active spectatorship. With respect to that, Peltonen et al. show, how digital installations in public space can create engagement and encourage passive bystanders to actively participate and interact with the installation and other users. Strongly connected to active spectatorship, Peltonen et al. further inform the **Interaction as a Performance** pattern [145]. They describe the use of content on the installation as well as interface features as resources to coordinate activity and create interaction events in order to give the interaction a performative character and making it meaningful for other people around.

A further behavioral pattern that frequently occurs with mutli-user interaction is **Conflict Management**. Conflicts usually range from ovelapping interaction to territorial conflicts in front or on the display or facade. The works by Peltonen et al. [145, 146] and Fischer et al. [55, 56, 54] showed that disagreements on shared resources do not necessarily produce a negative outcome. On the contrary, involving users and allowing them to solve conflicts on their own can create social encounters where users start socializing.

Depending on the size of the display or media facade and the respective interaction space in front of the screen, the aforementioned patterns can be observed in different occurrences. Most of the patterns require users being aware of other users. While for public displays, users can usually see each other, this might not always be the case for media facades due to their enormous size.

Guidelines

Vande Moere et al. investigated the contextual characteristics of media architecture, which are parameters that impact the integration in the existing social fabric [129]. They provide a set of guidelines addressing the contextual characteristics of media architecture from a socio-demographic (*environment*), technical (*content*) and architectural (*carrier*) perspective. While they conclude that media architecture does not seem well prepared to adequately respond to changes in its context over time, they provide a set of guidelines that target all relevant stakeholders, ranging from architectural designers to content managers and public authorities, with an aim to improve media architecture's acceptance and credibility, towards its long term sustainability in our urban fabric [129]. To get a better understanding of interactivity with public installations, Diniz et al. follow a similar approach and incorporate the dimensions of media facades into their guidelines [50]. They discuss the emergence of and need for principles and guidelines for the design of interaction spaces based on media facades. They combine empirically observed features of existing projects and map these onto a systematic framework of feature domains with the goal to guide the design of the active and interactive installations for media facades [50].

Azad et al. focus on territorial aspects of interaction with a shared public display, where they investigate behaviors on and around large vertical displays during concurrent usage [2]. Using an observational field study, they identify fundamental patterns of how people use existing public displays: their orientation, positioning, group identification, and behavior within and between social groups just before, during, and just after usage [2]. They further divide the behavior into inter-group (movements between groups and the environment) or intra-group (movements within a group) behavior, which they compare on and around the display. They found that for group behavior *on* a shared vertical display, the on-screen territories are similar to those observed in collaborative tabletop scenarios [118, 162]. They further show how territories of groups working concurrently influence each other and how these territories change throughout the interaction.

2.2.2 Interactive Media Facade Installations

Researchers, artists and designers have been exploring interactivity for media facade installations in various different ways. On the one hand, they experimented with different input modalities and interaction techniques in a playful and artistic manner. On the other hand, they shared their valuable experiences and lessons learned as a solid ground for developing novel, universal interaction approaches for media facades. In the following, we review interactive media facade installations, based on the utilized modalities.

Whole Body Interaction

With the design intervention *Aarhus by Light*, Brynskov et al. created an interactive installation for the concert hall in Aarhus, Denmark [30] (see Figure 2.9). The installation was installed on the glass facade of the concert hall building, which was fitted with $180m^2$ media facade consisting of semi-transparent LED screens. The goal of this installation was to engage local citizens into new kinds of public behavior to explore the potential of digital media in urban life. Within three interactivity areas in front of the facade, which where marked with colored carpets, the silhouettes of people standing on he carpet were tracked. Their movements were mapped to playful creatures on the media facade in order to encourage a curious and playful investigation of the expression among the users. The movement of the virtual characters further served as visual feedback on the users' movements. As described in Section 2.2.1, Brynskov et al. [30] revealed valuable insights around the themes of *interaction patterns*, re-occurring patterns, *initiation*, how do people engage with the installation, *interaction style*, how do people interact, as well as *relation*, denoting social interaction patterns. The ability of this design intervention to *support* engagement was further investigated by Dalsgaard et al. in [44]. They describe

Figure 2.9: Left: The Aarhus by Light installation at the Concert Hall Aarhus [30], where the silhouettes of users are mapped onto the facade. Right: The Climate Wall at the Ridehuset Aarhus. People can grab and move speech bubbles with their tracked body movements [45].

how the presence of territorial issues lead to social activities around the installation, where people have social encounters to understand and solve the territorial issues.

In [60], Fritsch and Dalsgaard present the *Climate Wall*, an interactive design intervention utilizing the whole body interaction approach from the Aarhus by Light installation (see Figure 2.9). The Climate Wall was an installation at the Ridehuset (an historical building) in Aarhus. It was installed and running during the climate conference *Beyond Kyoto*[1]. The installation displayed generated, fragmented climate statements, giving passersby the opportunity to take part in the ongoing climate debate. People could grab and move around individual words with their body movement, forming new climate statements. While interacting, the users received visual feedback on their input by the movement of the controlled word. The focus of these works was not the actual interaction technique. As the key aspects from the Climate Wall and Aarhus by Light installations, the authors revealed their insights into how people interact with such novel urban interfaces and how they can be engaged to interact. They utilized the media facades to create engagement, where the particular interaction technique supported the playful character of the installation.

[1] http://klima.au.dk/dk/forside/konferencebeyondkyotoconferen/

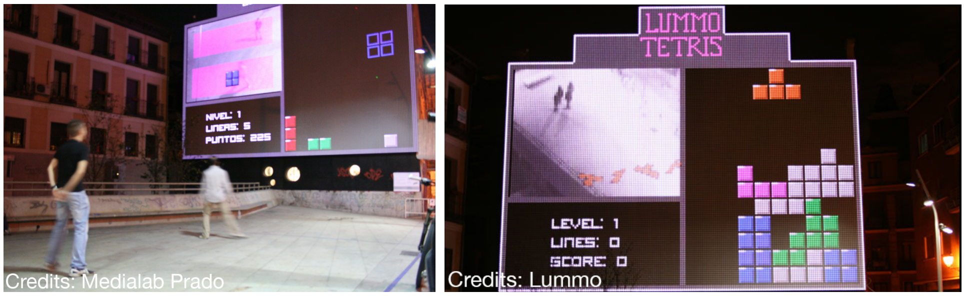


Figure 2.10: The Lummoblocks [80] installation at the MediaLab Prado, Madrid. Two player can play Tetris on the facade, while rotation of the blocks is controlled by the body movement of one player and moving the blocks by the body movement of the second player. Both within the boundaries of pre-defined interaction zones.

Gutiérez et al. show how whole body interaction can also explicitly enhance the experience of a user while interacting, taking the interaction to a spatial level. With Lummoblocks[2] [80], they created an interactive Tetris[3] game, running on the media facade of the Medialab Prado[4] in Madrid (see Figure 2.10). The aim of the installation is to provide a playful, interactive, space-located experience, as well as engaging social interaction between users and spectators. It was created in the context of the *Open Up* workshop in Madrid, Spain in 2010. The facade showed the visualization of the game, combined with a live video feed, displaying a birds-eye view of the users in the two interaction spaces in front of the facade as visual feedback. The goal of the game is to rotate and move the occurring blocks to create lines. The installation mapped both actions to two separate interaction spaces. Two players had to collaborate, where one had to rotate the block, while the second player moved it to the right spot. Both by running around within the boundaries of the particular interaction space.

With the installation *12m4s*, the LAb[au][5] group created an interactive media facade installation based on the speed an average person walks [103]. This architectural intervention uses the movements of passersby to generate a real-time visualization. They assume that a person walks an average distance of twelve meters in approximately four seconds. The movements of passersby are tracked in real-time with cameras to generate a visual (3D particles) and auditory (granular synthesis) scape on the facade, based on

[2]http://www.lummo.eu/lummotetris.html
[3]http://www.tetris.com/
[4]http://medialab-prado.es
[5]http://www.lab-au.com

the captured image data and Ultrasound sensors. The visualization is based on the position, orientation and speed of a passersby. It is projected on a 12m long Mylar[6] screen, combining projection and reflection of the passerby's body, while creating an interactive space in between the digital and the body space. The visual effects are amplified by the characteristics of the Mylar screen. The surface of the screen remains a mirror on non-enlightened zones and it turns transparent on enlightened zones.

The *Night Lights* installation by the YesYesNo[7] group combines three different interaction techniques [103] to display interactive silhouettes and animations on a projected media facade with a 3D projection mapping: tracking of body movements of people standing on a platform, hand gestures above a light-table and waving of mobile phones. The installation turned the Auckland Ferry Building in Auckland, New Zealand into an interactive playground. The goal of the installation was adapt the concept of *shadow puppets* to allow passersby to become performers. In front of the building, there are dedicated interaction zones for the particular interaction techniques, which are designed as small stages, where passersby can walk up and interact with the projection. The movements of people within these interaction zones are tracked and amplified by an approximately $15m$ tall projection on the wall of the building.

The building of the organization *La Vitrine Culturelle* in Montreal, Canada, is equipped with a small, low-resolution media facade of approximately $23m^2$ consisting of 35000 RGB LEDs that change their color as a reaction to the movements of passersby [82]. By connecting the interaction to the movements of passersby, the installation provides various animations and media content. When walking past the facade, the walking direction is mapped to animated arrows indicating the walking direction. When people stop and stand in front of the facade, they can create further animations by body gestures. These animations range from snowflakes popping up around the user's silhouette, to movable light spots. This installation was the first permanent interactive media facade in North America [82]. Initially installed as a temporary installation for Montreal's annual city of lights festival in 2009 as signage and branding for the building, it was turned into a permanent interactive installation due to its huge popularity.

[6] http://www.mueller-ahlhorn.com/en/mylar/
[7] http://yesyesno.com/

As the aforementioned installations show, interaction by body movement usually has a playful character and it comes with certain performative elements. Hence, whole body interaction is well suited for playful and performative installations, but its applicability as a general interaction technique to interact with complex content on distant screens is very limited.

Pointing

A common interaction technique that was utilized in various interactive installations is direct pointing. With *Spread.gun*, as a part of the VR/Urban group[8], Fischer et al. present an interactive shared encounter for media facades, inspired by established forms of graffiti culture [56]. The installation took place within the scope of the 2008 Media Facade Festival in Berlin, Germany. The aim of the installation was to create a digitally augmented forum in public space and to *reclaim urban screens*, which according to Fischer et al. are dominated by commercial interests. The tangibility of the interaction is mentioned as the most important part of the design work. The stationary interaction device — called *Spread:gun* — is a model of an ancient cannon. A user stands behind the cannon and rotates it to aim at a particular point on the facade. Additionally, a user can enter a text message through a digital touch screen. While aiming at the facade, a virtual red crosshair is displayed as visual feedback. The position of the crosshair is calculated with data from two potentiometers that are integrated in the canon-like interface for the horizontal and vertical axis. By shooting the cannon, a color bag is virtually shot onto the projected facade. The color spots on the facade are displayed enclosing the text messages of the particular user. While in general providing the possibility of interacting by direct pointing, this approach clearly has its limitations. In terms of design and interaction, the interaction device is highly tailored to the purpose of the installation and environment around the particular facade. Again, they point out the effect of the social and spatial setting and describe how the location and the surroundings may drastically alter the context of the installation.

As a conceptual enhancement of Spread.gun, Fischer et al. presented the *SMSlingshot* [55]. Instead of using a stationary input device, they provided a mobile, custom built input device, based on the metaphor of a wooden slingshot (see Figure 2.11). The aim of the installation was to create a digital slingshot with which people can throw information onto

[8] http://www.vrurban.org

Figure 2.11: The SMSlingshot installation. Left: The media facade displaying user-generated content. Right: The SMSlingshot device [55]. A user can shoot a message onto a facade by aiming at a specific point with the slingshot.

public screens. The slingshot device is based on Arduino[9] and it is equipped with a ultra-high frequency radio unit, a small LCD display, a laser and a mobile phone keyboard, allowing a user to enter text messages. By aiming and shooting at a particular point on the media facade wit the slingshot, a user can shoot his entered message on the facade together with a virtual color bag, analog to the Spread.gun approach. Again, while aiming at the facade, a laser pointer mounted into the slingshot is activated, which is tracked on the facade with the help of a camera but which also provides the user with visual feedback on the pointing target. Since the interaction device in this installation is mobile and not directly connected to the system as the Spread.gun is, the messages are transferred to the computer, directing the projection over a wireless radio modem. Additionally, the entered text messages are twittered in real-time. In comparison to Spread.gun, SMSlingshot has the advantage of providing a portable and flexible interaction technique. With *TXTual Healing*[10], Notzold followed an approach very similar to SMSlingshot, applying the same interaction metaphor and feedback mechanisms utilizing the model of an assault rifle to fire text-based messages onto a projected media facade [103]. While SMSlingshot displays color spots with the messages, TXTual Healing displays the messages within speech bubbles in order to reflect that the messages are actual quotes from real people. In its technical realization, TXTual Healing also follows the same approach as SMSlingshot. The position of a laser mounted onto the rifle is tracked on the facade by a camera.

[9] http://www.arduino.cc/
[10] http://www.txtualhealing.com/

Figure 2.12: Left: The technical setup of the Laser Tag installation, comprising two projector for creating the media facade. Users are interacting by pointing at the facade with a laser pointer whose movement is tracked by a camera. Right: Content created by users.

For the described approaches, a user only points at the media facade to transfer locally created content onto the facade. In Contrast, the installation *Laser Tag* by the Graffiti Research Lab[11] allows for continuous pointing interaction to create content directly on the actual media facade [103]. As depicted in Figure 2.12, users can paint and write onto the media facade by pointing at it with a laser pointer. The position of the pointer on the facade is tracked with a camera and the movement of the pointer is virtually repainted on the facade using a projector, which results in immediate visual feedback on the pointing. The scope of this installation was to create a playful approach for allowing people to digitally write on buildings in public spaces.

The aforementioned installations all have in common that they utilize a pointing approach where a laser is tracked by cameras. Although this is a trivial, from a technical perspective easy to realize approach providing immediate visual feedback on the pointing, it's applicability as a general interaction technique for permanent media facades that are not permanently monitored is rather limited. From the technical side, the tracking does not work with standard, free to use laser pointers. The utilized lasers require special permissions to be used in public. From an interaction perspective, it is barely possible to distinguish between users or limit the interaction to particular users since there is no connection between user and laser. Such a system only detects a certain number of laser

[11] http://www.graffitiresearchlab.com

Figure 2.13: Left: The technical setup of the Wall of Light installation. A user is painting on the facade using the Tag tool. Right: Content created by users.

points on the screen but it is not aware from how many lasers they are coming. However, interacting with a media facade by pointing at it revealed as an intuitive and promising approach which we will further investigate within this thesis.

Touch & Gestures

A further approach explored in different media facade installations is allowing input by direct touch on a digital surface or by gestures performed with an input device such as a mobile phone.

Within the scope of the Illuminating York Festival[12], in conjunction with the GaiaNova Productions Ltd.[13], the OMA International group presented the *Wall of Light* installation [103], an artistic installation to paint in real-time on a projected media facade. As illustrated in Figure 2.13, using the *Tag Tool*[14] — a digital graphics tablet for painting — as the input device, people were able to paint onto the facade by drawing on the Tag Tool with a digital pen. The Tag Tool itself does not provide any feedback on the input locally and hence, the content appearing on the facade serves as visual feedback on the users' drawings. The installation was intended as a canvas for daily performances of artists,

[12]http://www.illuminatingyork.org
[13]http://www.gaianova.co.uk
[14]http://www.omai.at/

but they also allowed visitors to use the installation to express themselves. According to the creators, the performance of the inexperienced visitors shows that the installation is intuitive and easy to use.

The urban media installation *Touch* by the LAb[au] group allows passersby to interact with the Dexia Tower[15] in Brussels, Belgium [103]. They installed a stationary multi-touch screen at the base of the building. With this multi-touch screen, people can individually or collectively manipulate the color of the approximately 4200 windows of the 145m high tower in real time. This window animation facade allows for both dynamic and static inputs. People can create basic geometric forms on the tower as well as change its color. The concept of this installation tries not to treat the facade as a plain 2D surface. The goal is to integrate characteristics of the building, such as orientation, scale or volume, to create an interactive experience. The stationary screen displays a virtual representation of the buildings surface, on which the users get immediate visual feedback on their interactions.

With *MobiSpray* [157], Scheible and Ojala presented a digital graffiti tool utilizing a mobile phone as a virtual spray can (see Figure 2.14). Users can spray digital dabs onto a projected media facade. The movement of the spraying nozzle is based on the orientation of the mobile phone, which is determined by continuous data from the built in accelerometer. The sensor data is mapped onto x- and y-coordinates on the facade. As can be seen in Figure 2.14, a drawing client application on the mobile phone provides controls to adjust the properties of the virtual spray can such as blob size, colors or intensity. A user can draw onto a facade by performing spray gestures with the mobile phone, similar to a real spray can. While feedback on the moving direction and the settings of the virtual spray can (e.g., color, size of the spraying nozzle) is displayed locally on the mobile device's screen, visual feedback on the painting is show immediately on the facade itself. Since there is no absolute mapping between the spray canvas and the facade, the advantage of this approach is that it does not require any calibration when used in a different location for a different building.

[15]http://www.emporis.de/building/dexiatower-brussels-belgium

Figure 2.14: Left: The mobile interface of the MobiSpray installation. Users can change the paint parameters with the key pad [157]. Right: A user painting on a public building using MobiSpray [157].

Shamma et al. utilized and adapted MobiSpray to create a collaborative space for the voice of a local community to make statements about the world around them [165]. Participants and spectators can observe how digital graffiti is created while they receive feedback provided by dancers interacting with the virtual projected content and physical space. To create graffiti, people can use MobiSpray and they can additionally upload text messages as well as images from the mobile phone to the media facade.

Besides MobiSpray, the presented installations are both stationary and highly tailored to the one specific setting and usage scenario. This limits their applicability as general input techniques to interact with media facades. With MobiSpray, Scheible and Ojala have given an example of an interaction approach that is (1) easily portable to different locations and differently shaped facades and (2) does not require calibration of the interaction device and the screen. However, MobiSpray requires a dedicated mobile phone as the input device and hence limits the number of users to one. Furthermore, it is specifically designed for painting on a projected facade and its applicability for different settings and tasks is limited.

Remote

Within the scope of this thesis, we are investigating *direct* interaction with media facades in a sense that to be able to interact with a facade, the user has to be within the potential interaction space in front of the facade. Especially for high resolution facades displaying advertisements, operators can remotely interact with the facade over web interfaces or dedicated software in order to provide the displayed content. Although they are interacting with the facade, we do not address this form of interaction since it can be instead considered as pure provisioning of content and it does provide passersby with the opportunity to interact with the facade. However, there exist different approaches in which people have been provided with a remote interface to interact with a media facade without the requirement of being within the potential interaction space or even the display space. While the aforementioned installations focused on allowing the people within the display space to interact with the facade, providing remote interfaces is usually driven by different design intentions. First, providing remote interfaces does not particularly focus on the people within the display space. It allows a wider range of people to interact with a facade while the interaction could be either limited to a particular group of people like for example all inhabitants of the city, or it could be allowed to all people that can access the remote interface (e.g., over the internet). Second, media facade installations providing remote interfaces often do not focus on continuous interaction. They rather provide means for sharing content with a media facade (e.g., by uploading images, videos or animations) or to participate by expressing yourself and visually branding or customizing a building.

A well known and one of the first media facade installations is the *Blinkenlights* project in Berlin, Germany, realized by the Chaos Computer Club[16] [81]. Within the scope of the 20th anniversary of the Chaos Computer Club in 2001, the upper eight floors of an office building were turned into the at that time world's biggest interactive computer screen. As depicted in Figure 2.15, they created a window raster animation facade by equipping the windows of the building with 144 individually controllable lamps in total, which resulted in display with $8x18$ pixels. To control the content, people had to call a dedicated phone number with their mobile phones. Once connected, they could either control a virtual cursor on the facade or activate previously uploaded animation with a

[16] http://www.ccc.de

Figure 2.15: Left: The Blinkenlights installation in Berlin, Germany [81]. An office building was turned into a window raster animation facade by equipping its windows with 144 individually controllable lamps. Middle: The Marnix media facade in Brussels, Belgium. People can download an animation interface to remotely control the displayed content [82]. Right: The Rundle Lantern media facade in Adelaide, Australia. People can create and upload their own content through a provided web interface [82].

code also entered with the key pad of the mobile phone.

The world's first interactive RGB media facade, the *Marnix* facade in Brussels, Belgium, allows people to participate by taking control of the media facade's content [82] (see Figure 2.15). The operators of the facade provide a free-to-download animation interface, with which people can create and upload their own animations from anywhere in the world. Whenever their animations appear on the facade, the creator of the animation is notified by email thanking them for taking part. This installation is very popular and frequently used from people around the world to share content from wedding proposals, to political statements, to artful animations.

The media facade of the *Rundle Lantern* building in Adelaide, Australia was built with *"the vision of creating an experience that would capture the imagination of the city and add beauty to people's lives"* [82]. As depicted in Figure 2.15, this low-resolution media facade consists of more than 700 individually controllable RGB panels that are wrapped around the building. According to its creators, the facade is intended to be a platform where people can express themselves, rather than a fixed interaction concept. To control the content of the facade, people can use a web interface to create and upload their own animations. With its frequent use, this installation has developed into a tool for continuous community engagement.

The interactive media facade installations reviewed in this section all have a playful, performative or artistic orientation. The interaction as well as the applied input techniques and devices were particularly designed for the specific installation. To establish media facades as future urban computing surfaces, we need generalized interaction techniques, which can be easily used in various settings with different properties. Besides the playful and artistic character of the installation, one important aspect when interacting with a media facade is its resolution and form factor. As we have seen, the resolution can easily vary from screens with 144 pixels to high-resolution screens capable of displaying sophisticated content. Furthermore, the majority of the presented installations require a direct line of sight between user and facade to interact. However, due to their size and form factor, not all parts of a facade might be visible at all times. This would leave a user with only a part of the actual facade as an interactive screen. When developing novel interaction techniques, we need to address these issues in order to exploit the media facades' full potential as urban screens.

2.2.3 Interaction with Public Displays

When interacting at a distance with a media facade, a user is generally interacting with a remote display in a large scale setting. The research on interaction with situated public displays or remote displays in general usually deals with interaction at-a-distance in a comparably small and controlled setting. This allows for instrumenting the display as well as the potentially interactive area around it, in order to realize novel and more complex interacting techniques. Early work by Rekimoto investigated the potential of interacting with distant screens in continuous work spaces [151]. However, due to the size of media facades and their large scale, uncontrolled and dynamic environment, the applicability of such techniques for media facades is usually limited. In this section, we therefore give a brief overview on existing work on interaction with distant displays grouped according to the utilized input mechanisms.

Whole Body Interaction

With [19], Bolt provided a first approach to interact with a distant display using speech and gestural input. The system allows a user to create different geometrical shapes at arbitrary locations on a projected screen in an instrumented media room. With commands like "*Put a red circle there!*" while pointing at the target location, the objects are created at the particular location. Ballendat et al. introduce the notion of *proxemic interaction*, which they consider as *devices with fine-grained knowledge of nearby people and other devices — their position, identity, movement, and orientation — and how such knowledge can be exploited to design interaction techniques* [6]. For the use case of an instrumented multimedia environment, they demonstrate how proximity and orientation of people and devices can be exploited to control multimedia content on a remote screen. *Light Space* is small multi-display environment installation, equipped with multiple depth cameras and projectors mounted to the roof of the room [198]. Wilson and Benko describe how within LightSpace (compare Figure 2.16), selective projection of data from the depth cameras can enable the emulation of interactive displays on un-instrumented surfaces (e.g., tables or desks), as well as it can enable mid-air interactions between and around the displays [198]. In addition to interacting with one of the surfaces, users can transfer content in between the interactive surfaces as depicted in Figure 2.16 by either simultaneously touching a source object on one screen and the target screen or by swiping an object from one of the surface onto the user's hand then dropping it onto another surface by touching the surface. Beyer et al. utilized multiple Microsoft Kinect[17] depth cameras to provide an interactive, cylindrical public display, allowing passersby to control virtual avatars by whole body gestures [15]. When we recognize a public display, we naturally use our eyes to engage with it. Exploiting this, *SideWays* provides a person-independent gaze interface for spontaneous interaction with a distant display, with which users can scroll through multimedia content without any prior calibration [205].

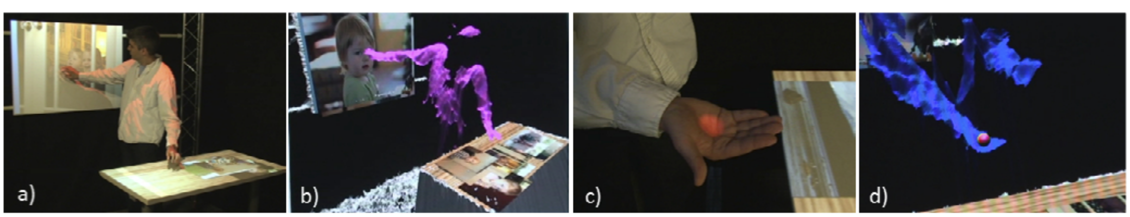


Figure 2.16: The LightSpace prototype combining depth cameras and projectors to provide interactivity on and between surfaces, including object transitions through-body between existing interactive surfaces (a-b), as well as interactions with an object in hand (c-d) [198].

[17]http://www.microsoft.com/en-us/kinectforwindows/

Specialized Input Devices

A common approach to interact with distant displays is the use of specialized input devices. Baudisch et al. introduced *Soap*, a mid-air pointing device consisting of an optical sensor device moving freely inside a hull made of fabric [10]. The device's core component is depicted in Figure 2.17. Kristein and Müller describe how to use a camera-tracked laser pointer to interact with a projected screen [102]. Olsen and Nielsen follow a very similar approach, providing calibration techniques for mapping the laser pointer to display coordinates [139]. Vogt et al. demonstrated how to realize collaborative multi-user scenarios using different laser pointers [188]. To allow 3D gestures in addition to controlling a pointer, the *XWand* is equipped with accelerometers and magnetometers [197] (see Figure 2.17). XWand is a wireless pointing device to allow styles of natural interaction in instrumented environments. While the aforementioned approaches are single-user scenarios, Cheng and Davis presented *LumiPoint*, a system allowing multi-user interaction with large distant screens using laser pointers [48]. Myers et al. [133] and Banerjee et al. [7] both compared pointing with a laser pointer to interact with a distant screen to manual pointing, as well as differently shaped pointing devices.

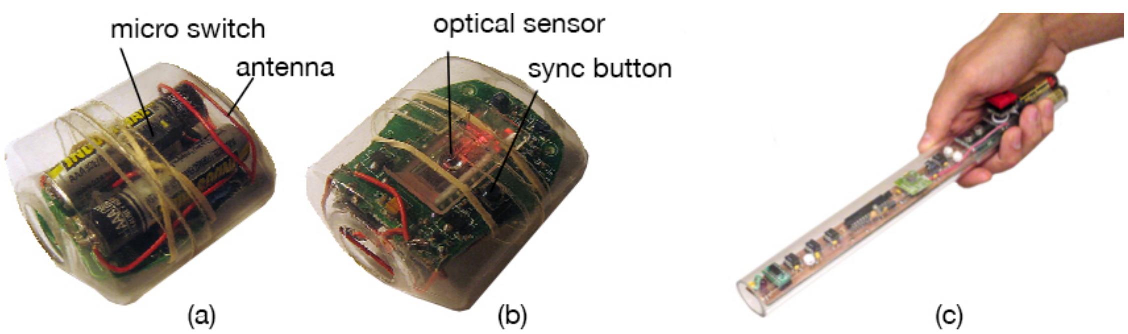


Figure 2.17: (a),(b): Top and bottom view on the core component of Soap [10]. (c): The XWand device [197].

Due to the natural hand tremor, manual pointing with a laser pointer is error-prone. To address this, Myers et al. introduced *Semantic Snarfing*, and interaction technique where a laser pointer is mounted onto a mobile device to indicate an area of interest on the remote display [132]. The selected area is then displayed on the mobile devices screen to allow for more precise input by manual touch. With *Shadow Reaching*, Shoemaker et al. introduced a technique utilizing a perspective projection applied to a shadow representation of a user, where a user can reach distant content by casting a shadow [168]. The applicability of specialized input devices to interact with media facades is determined by particular use case and the purpose of the installation. While they could limit the generalization and transferability of the installation, they can also convey a metaphor as in the case of SMSlingshot [55], being a dedicated part of the overall installation.

Mobile Devices as Input Devices

Over the last decade, with the involvement of mobile phones from pure communication devices to powerful, smart computing devices equipped with various sensors, cameras and a constant wireless data connection, mobile devices have gained importance as input devices to interact with distant screens. Ballagas et al. surveyed different interaction techniques utilizing the mobile device as a ubiquitous input device to various resources available such as situated displays, vending machines and home appliances [4]. They argue that due to their ubiquity, mobile phones have great potential to become the default interface to ubiquitous computing environments. Paek et al. also aimed to turn mobile phones into universal interaction devices for public displays [140]. They created an interaction platform, allowing users to use their personal mobile devices as input devices for public displays. Cheverst et al. further explored mobile phone interaction with small situated public displays, such as digital door signs [41]. As we can see in a wide range of research, due to their technical capabilities and their permanent availability, mobile devices indeed turned out to be the most dominant external input device for interacting with distant displays. In this context, mobile devices are used for transferring user input to the distant display, as well as for redirecting content from the distant display to the mobile device, where it is manipulated locally. Furthermore, mobile devices can be used as a *second screen*, a mobile, private extension to the distant public display. Within the scope of this work, we are particularly interested in the use of mobile devices for transferring user input to a distant screen. Hence, we focus on the corresponding existing work. When using a mobile device as an input device, the interaction itself is carried out in various ways. Rukzio et al. compared interaction with a distant display utilizing a mobile device by *touching* the display with the mobile device, *scanning* [154] based on Bluetooth[18] and pointing at the display. In particular, they analyzed when a given mobile interaction technique should be used as a selection technique in smart environments. They

[18] http://www.bluetooth.com

found that people prefer to touch things that are close. If they're not close and there's a clear line of sight, they prefer pointing [154]. Next, we review related work that utilizes mobile devices to interact with remote displays in different shapes.

Touching the display

When interacting with a remote display, the most direct way would be to touch the display directly. While it is the standard approach to interact with touch capable displays such as the *DiamondTouch* [49] or common interactive public displays such as the *UBI Hotspot* [138], there is an increasing number of approaches utilizing touching the display with a mobile device. For static displays, such as maps and posters, there exists a number of approaches utilizing *Near Field Communication*[19](NFC). In this case, the static displays are augmented with NFC tags providing information that can be read by touching them with an NFC-capable mobile device [149, 189, 87]. To interact with a dynamic projected screen, Hardy and Rukzio mounted a grid of NFC tags behind a projection screen [86, 88]. Users could select a certain area on the display by touching the screen at the particular location on the screen with a NFC capable phone. Broll et al. adapted this approach to create a Whack-a-Mole game for dynamic public displays to investigate how people appreciate using NFC tags [29]. With *PhoneTouch*, Schmidt et al. further refined this technique enabling users to select targets on a tabletop by direct touch with a mobile device [159]. As depicted in Figure 2.18, when touching the digital surface with the mobile device, the touch point detected by the surface and the device identity provided by the mobile device based data from the built-in accelerometer are merged based on their correlation in time. Schmidt et al. further provided a characterization of interaction with a digital surface by touching it with a mobile phone in terms of input, output and contextual attributes [160]. They analyze PhoneTouch with respect to its suitability for transferring data, personalizing content and user interfaces, authenticating users, providing personal localized feedback, as well as the expressiveness of input using PhoneTouch.

[19]www.nfc-forum.org

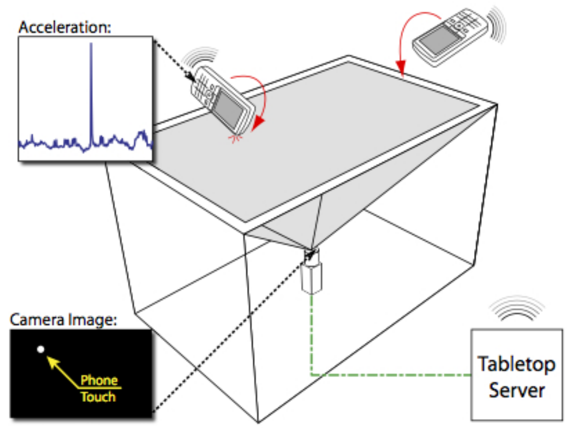


Figure 2.18: Touching a surface with mobile phone using PhoneTouch: Surface and phones independently detect the touch events and communicate the observations over a wireless network, correlated in time, and combined to associate a touch with both surface position and phone identity [159].

The applicability of touching the screen for interacting is very limited within the scope of media facade. Due to their size and the required minimal viewing distance, users (1) usually interact from a distance where they facade cannot be touched directly and (2) if they would approach the facade to touch it, they could only interact with a fraction of the overall facade. As a result, touching a media facade — reading NFC tags — could only be an option to connect a mobile phone to a facade or to transfer information to the mobile device, but not for continuous interaction.

Sensor-based input

Mobile phones nowadays are equipped with a variety of sensors, including sensors such as accelerometers, gyroscopes and electronic compasses, which can provide information on the orientation, movement and acceleration of the mobile device within a reference space (i.e., a local coordinate system). The data provided by these sensors can be used to interact with external devices, such as digital screens. This approach is also utilized in commercial products, such as the Nintendo Wii[20] gaming console, which comes with a remote controller to steer a pointer on the TV screen based on orientation and acceleration of the controller. However, accelerometer-based input is still also a widespread topic in research.

[20]http://www.nintendo.com/wii

Picking up on this, Vajk et al. describe how to turn a mobile device into a Wii-like game controller [186]. Rekimoto initially presented an approach using the tilt of a mobile device to move around on a digital screen [150]. With *Rock'n'Scroll*, Bartlett follows a similar approach using inertial sensors on a mobile device to allow for performing gestures with a mobile device to scroll through a photo library and select photos [9]. The same approach was investigated from a rather technical perspective by Cho et al. [42]. *Toss-It*, a system presented by Yatani et al., allows users to transfer images to a distant display by performing a throwing gesture [204]. Tilting a mobile device has also been applied by Dachselt and Buchholz to transfer images to a distant display and to scroll through the available content [43]. This approach is similar to *MobiToss*, presented by Scheible et al. [158]. MobiToss allows for transferring multimedia art — including videos — from a mobile device to a remote display and for manipulating the transferred content on the display.

Camera-based interaction

A further widespread approach for using mobile devices to interact with a digital display is camera-based interaction, where the discrete or continuous images from the mobile device's built-in camera are used to compute the interaction parameters. The interaction itself can happen in various ways from controlling the pointer, to connecting and accessing digital services. Toye et al. studied using camera equipped mobile phones and visual tags to access mobile services [181], finding that novice user can easily handle selection tasks using visual tags.

With *Sweep* and *Point & Shoot*, Ballagas et al. describe two techniques allowing users to interact with a distant screen based on recognition of visual tags with the mobile device's camera [5]. It allows both selecting targets on the remote screen, as well as using the mobile device in a mouse-like way with multiple degrees of freedom to control a pointer [5] (see Figure 2.19). Jiang et al. present a more general version of this approach, investigating the performance when utilizing a mobile hand-held camera to continuously control a pointer on a remote display [99]. The term *mixed interaction space* was introduced by Hansen et al., arguing that the possibility of using mixed interaction spaces is what distinguishes camera-based interaction from other types of sensor-based interaction on mobile devices. [85, 84]. They used computer vision to calculate the distance between mobile device and remote screen and they adjust the zoom level of a displayed image based on that distance (see Figure 2.19). With *Scroll, Tilt or Move It*, Boring et al. propose and compare three techniques to continuously control a pointer on a remote screen, in order to allow for bidirectional interaction [25]. The movements of the mobile device are mapped to the pointer movement on the remote display, allowing the user to also control the speed of the pointer tilting the mobile device.

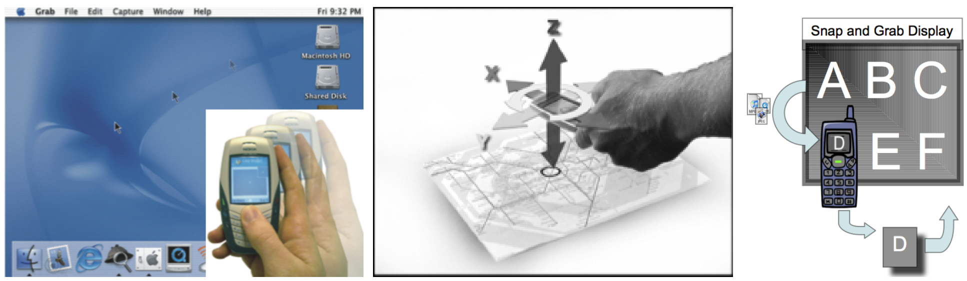


Figure 2.19: Left: Controlling a pointer with the *Sweep* technique [5]. Middle: Calculating a mobile device's position and orientation to a remote display with 6 degrees of freedom using computer vision [85, 84]. Right: Transferring multimedia content over Bluetooth applying the *SnapAndGrab* technique [120, 119].

In contrast to controlling a pointer on a remote display, researchers also used camera-equipped mobile phones to visual transfer information from a display to a mobile device. With *SnapAndGrab*, Maunder et al. provide a technique allowing users to access and share multimedia content with situated public displays and their own Bluetooth-enabled camera phone [120, 119]. As depicted in Figure 2.19, to select content, they take a photo of a visual download key shown on the remote display. After processing the key, the selected content is transferred via Bluetooth to the user's mobile device. Taking this approach one step further, Boring et al. present *Shoot & Copy* [21]. The system allows to shoot a photo of any part of a public display with their mobile device. It automatically processes the image to determine the corresponding part of the display on the photo, in order to immediately retrieve the displayed content to the mobile device over a wireless connection. In contrast to SnapAndGrab, Shoot & Copy does not rely on visual codes. The same approach was followed by Chang and Li presenting *Deep Shot*, a framework for migrating a wide range of common tasks (i.e., viewing web pages, editing calendars, etc.) across devices [40]. In particular, Deep Shot supports migrating tasks in both directions, from a remote display to a mobile device and from a mobile device to a remote display.

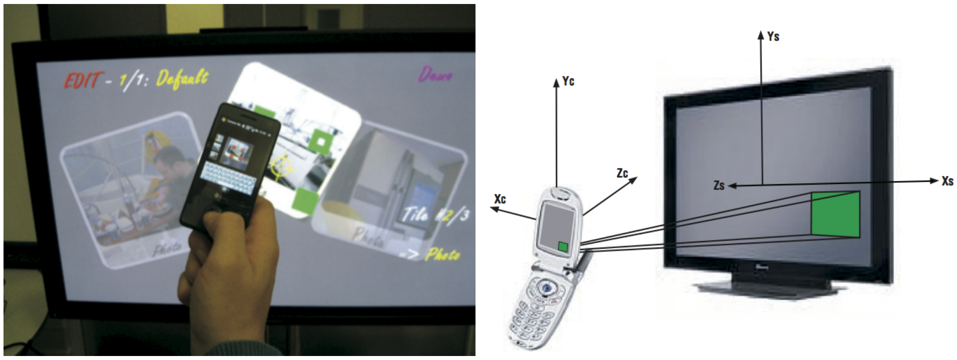


Figure 2.20: Left: Images being moved, rotated and scaled using the display registration by Pears et al. [144]. Dynamic visual markers are displayed around the particular content elements. Right: The geometry of a smart phone viewing a marker on a public display [144].

Tani et al. proposed to remotely operate machines in a factory by manipulating live video on a remote computer screen to envision how users could interact with distant real-world objects [179]. This approach was adapted in different works to utilize mobile devices as a see-through interface or *Magic Lens* [16, 17]. Pears et al. adapted this approach to interact with a distant digital display through live-video on a mobile device by creating a geometric mapping — called *display registration* — between the visual representation of the display in the live-video stream on the mobile device and the distant display itself [144] (compare Figure 2.20). To create this mapping, they utilize perspective distortion, for which they initially expect the remote display to be planar and rectangular shaped. They display visual markers around particular content elements, which are detected in the live-video on the mobile device using computer vision techniques. To calculate the perspective distortion, top- and baseline of the marker are aligned against each other. By registering both displays, they use the mobile device as a 6-degree-of-freedom flying mouse (compare Figure 2.20). With the registration, each pixel on the distant display is mapped onto a corresponding pixel on the live video displayed on the mobile device's screen. This further allows direct interaction with elements on the remote display by direct touch input on its visual representation on the mobile device's screen. Direct touch on the smart phone's image of the remote display is indistinguishable from directly interacting with the remote display itself [144].

Boring et al. further extended this approach with *Touch Projector* [23]. Similar to the the display registration technique by Pears et al. [144], Touch Projector enables interaction with remote screens through a live video image on a mobile device where they *project*

Figure 2.21: Moving content between displays using *Touch Projector* [23]: (a) the user aims at a display (b) causing the device to automatically zoom in. (c) This allows the user to touch the red target object and (d) hold it while turning the device towards the other display. (e) Once the device detects the secondary display, (f) it zooms in again. (g) Pushing the freeze button causes the live camera image to pause for precise manipulation. (h) Lifting the finger releases the object [23].

the touch input to the remote display. As a conceptual enhancement, Touch Projector provides a system architecture, allowing for cross-display interaction in *multi-display environments* [3]. As depicted in Figure 2.21, in addition to manipulate content on arbitrary screens within the environment, users can also select and drag content between displays. When observing the remote display through live video from different distances, the remote display will appear in different sizes in the live video, depending on the distance. When interacting with the remote display through live video on the mobile device, this results in different sizes of the interaction canvas, influencing the accuracy of the touch input. To overcome this, Touch Projector provides an automatic zoom functionality. When detecting a display, the system automatically zooms in to a constant aspect ratio. Since zooming amplifies the natural hand tremor resulting in jitter, Touch Projector further provide a freeze mode, in which the user can temporarily freeze the current view to obtain a stable image for the interaction. With *Virtual Projection* [12], Baur et al. address the problem of limited screen real estate for techniques like Touch Projector. They apply the metaphor of optical projection to digital surfaces in a multi-display environment, inspired by how people intuitively control an image's position, size and orientation with a hand-held projector. Figure 2.22 illustrates how a user can project an image onto a display and how he can manipulate and move it. Using optical tracking of image features, Virtual Projection allows selecting a target display by a single gesture, i.e., pointing at it with the mobile device's camera. The orientation of the mobile device to the display is determined with computer vision algorithms based on the observed perspective distortion of the remote display with the mobile device's live video. Manipulating images on the remote display is based on gestures consisting of touch input on the mobile device's screen in combination with the movement of the device. Boring and Bauer further combined both approaches into a conceptual framework for making distant public displays interactive [22].

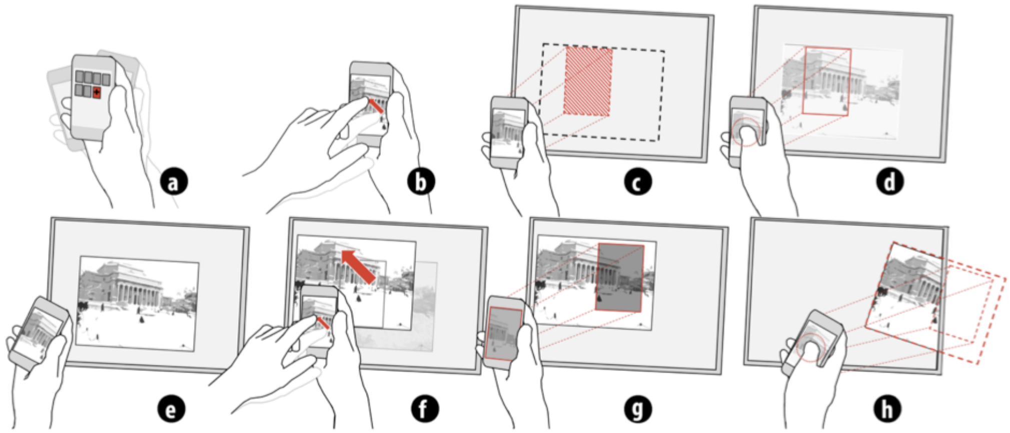


Figure 2.22: Virtual Projection [12]: (a) Shaking the device to create a view. (b) Interacting with a non-projected view. (c-e) Creating a projection by aiming at the secondary display, pressing and holding, and releasing. (f) Synchronized interaction. (g) Projection frustum can be used for filtering or navigating. (h) Projections can be moved or deleted by aiming, pressing and holding, and dragging out of the display [12].

From the aforementioned approaches, we see two major advantage of interacting through live video on mobile devices, making it promising candidate for interacting with media facades in an urban context:

1. The presented approaches require a comparably low instrumentation of the interactive environment. The computation is solely based on visual means and it only requires a constant wireless data connection between mobile device and remote display.

2. Interaction through live video provides an intuitive way of interacting with remote objects and displays by direct touch.

However, these advantages can easily turn into disadvantages. Changing weather and lighting conditions can drastically decrease the accuracy of the optical tracking. Furthermore, the resolution of the media facade can also limit the accuracy of the mapping of touch input on the mobile device to the media facade. Hence, interaction through live video on mobile devices needs further research to make it applicable in the context of media facades.

2.3 Interaction Design and Development

2.3.1 Designing Interaction for Media Facades

When designing interactive media facade installations, the unique properties and requirements of this whole domain need to be considered in order to create successful and enjoyable installations. First explorations towards designing such installation on a large-scale for public settings have been made by Fritsch and Dalsgaard [60]. While analyzing their two aforementioned installations *Aarhus by Light* [45] and *Climate Wall* [30], they identified affective experiences and engaging interaction as the two main perspectives in order to provide an interactive experience for the long run.

One of the first systematic analyses of the design of interactivity for media facades was presented by Dalsgaard et al. [47]. Providing the *Design Space Explorer*, they present a framework for managing multiple sources of information and domain concerns for the design of collaborative design process, which combines these aspects in scenarios for design concepts [47]. The Design Space Explorer captures and gives an overview of design materials and forms, domain locations and situations, interaction styles, and content types. It is a tool to gain an overview of the design space, to conceptualize key aspects of interaction design, as well as to support communication and discussions among designers, clients and partners. Furthermore, it provides a platform for designers to combine these aspects into scenarios for design concepts [47]. Dalsgaard and Halskov further expanded this work [45], identifying eight key challenges that need to be addressed when designing for media facades in an urban context. These challenges consider a wide range of issues:

1. As already mentioned, existing interfaces cannot be transferred without further effort. Urban settings call for new or adapted forms of interfaces.

2. New installations need to be integrated into existing surroundings.

3. Changing light and weather conditions must be taken into account.

4. The content has to suit the medium. It has to match the technical properties of the facade and it needs to support the intended interactions.

5. Stakeholder interests need to be balanced. This can be a critical issue, since the majority of media facades are owned by companies or public institutions enforcing strict rules about their presence in public.

6. The diversity of situations in the highly dynamic urban space needs to be considered.
7. The introduction of new technologies might influence social behavior.
8. Technology might be used in a different way than intended, even misused.

The latter is controversial argument. Eriksson et al. argue that any part of public space can be misused and introducing a new channel only makes it different, not better or worse [53]. Furthermore, they claim that being able to provoke — and hence to inspire — is a basic characteristic of public space and they identify regulation as an important design aspect. Dalsgaard and Halskov further point out the importance of documenting the design process to allow for reflective design, learning from past experiences [46]. Vande Moere and Hill analyze the concept of research through design in the context of teaching urban computing [128]. They further point out the importance of the context of an urban installation for its design process [129]. Reeves et al. provided a taxonomy to classify public interfaces according to the extent to which the interactions of a user and the resulting effects are hidden, revealed or amplified for spectators [148]. Korsgaard et al. approached the question of how to design for media facades by discussing how they structure their design processes to address specific sets of challenges raised in the literature [105].

2.3.2 User Experience

Prototyping User Experience

A systematic design process allows for developing successful applications [34, 166]. Moggridge [130] argues that the core of interaction design is to create interactive experience prototypes [31] while gathering continuous user feedback. Testing prototypes of the design concept in early stages allows for more design iterations and increases the quality of the outcome. This is referred by Buxton as getting the design right before getting the right design [34]. During the design process, usually both low-fidelity and high-fidelity experience prototypes are created [153]. They depend on (1) how accurately they need to represent their real-world counterpart, and (2) the number of details and/or functionality they have to include. They closely correlate to the aim designers pursue with a prototype, which is commonly referred to as the prototype's scope [115].
When designing interactive systems for media facades, designers should aim for iterative testing design components dealing with media facade interaction as early as possible.

Methods for creating such experience prototypes [31] include, for example, Wizard-of-Oz techniques [34], acting out scenarios, or electronic prototyping platforms [77, 122]. Such tools help designers to create prototypes early in the design process in a more time- and cost-effective manner. When designing for media facades, however, there is no common ground for creating experience prototypes.

Evaluating Experiences

Evaluating interaction with media facades can be complex in many facets. There is a fluctuating audience and a large number of users, as well as rapidly changing weather conditions. Since media facades are usually not active during daytime and they are permanently occupied, time slots for testing and evaluating are limited. Furthermore, multiple users can interact simultaneously with the facade and with each other through the facade without seeing each other. A user's reason to interact with an interactive installation might be to achieve a particular goal, but instead the experience of the interaction itself. Hence, we need to consider methods from UX approaches to evaluate media facade installations.

An important aspect when evaluating UX is to understand the user. Wright and McCarthy review emerging design and UX methodologies in terms of dynamically shifting relationships between designers, users, and artifacts [202]. They outline that if experience is central to designer-user relationships, emphatic methods have to be understood and used in an appropriate way. Furthermore, they describe their understanding of experience-centered design as a humanistic approach to designing digital technologies and media that enhance experiences [203]. Forlizzi and Battarbee addressed the diversity of experience for interactive systems, where they characterize existing approaches to experiences and provide a framework for designing experiences originating though interactive systems [58] . Furthermore, they argue that for novel technologies, an experience-oriented design approach is the only way that user-centered design can have a valuable impact on the design.

Bargas-Avila et al. review different methods for designing and evaluating UX [8]. They demonstrated that UX methods refer to emotional aspects of an experience when interacting with a system. To evaluate such matters, Hassenzahl et al. developed AttrakDiff, a scientifically-applicable tool for measuring the pragmatic quality, attractiveness, identity, and stimulation of the interaction with a product or service [91]. The positive and negative affect schedule (PANAS) [190] measures and explains positive and negative effects of a retrospective experience. Creating experience diaries where participants write a report about their use of a product over the course of weeks [104, 141] is another common approach. Geven et al.[73] proposed storytelling as a tool to evaluate user experience in narrative interviews [141]. In the form of experience reports, storytelling is also used to collect information on meaningful experiences with interactive products or services [104]. However, triggering participants to state their experiences in the interview sessions was accomplished by making reference to familiar electronic products. Burmester et al. described the valence method [33], an approach that evaluates the emotional quality of an interaction in two phases: (1) In the formative phase, the user records positive and negative feelings while interacting with a product or service. (2) In the summation phase, the interviewer asks participants about reasons for their actions during the interaction, using an in-depth interview method [152], until they can be matched to the underlying psychological need [167]. This model is also based on Hassenzahl et al.'s UX model [92]. It reduces the complexity of UX with the help of positive psychological needs (such as the feeling of autonomy and competence). These practices originate from a different context; we need to adapt existing approaches to the context of media facades, where users interact with a novel system in limited time-slots.

2.3.3 Simulation & Prototyping

With the increasing availability of 3D printers, creating detailed small-scale models, e.g. to represent a building and its surroundings, is technically feasible in no time. Jacobs introduces *rapid prototyping* and describes the level of accuracy that can be reached, providing large benefits to the designer at low cost [96].

In [134], Nakanishi shows that physical and digital representations can also be combined when prototyping public display applications. However, he remarks that this approach comes with the supplementary issue of synchronizing the virtual and physical parts of the toolkit. Nakanishi et al. further provide a framework for hybrid prototyping, called the *City Compiler* [135]. The framework aims at bridging the gap between the two realms (see Figure 2.23). Providing the designer with an impression of the application's performance in a virtual environment and allowing "interactive trial-and-error" testing, the City Compiler supports an iterative design process. However, the main focus of city compiler is to provide a toolkit to prototype the visual appearance of an installation in virtual and hybrid simulations. This addresses use cases like digital signage in retail stores or also in urban public spaces. City Compiler does not sufficiently address interactivity in the design process. Additionally, current tools to completely execute an application on a simulated facade such as the City Compiler are still very limited. They are tailored to one specific facade and they neither support interactivity, nor do they provide opportunities to model and simulate the surroundings of the media facade.

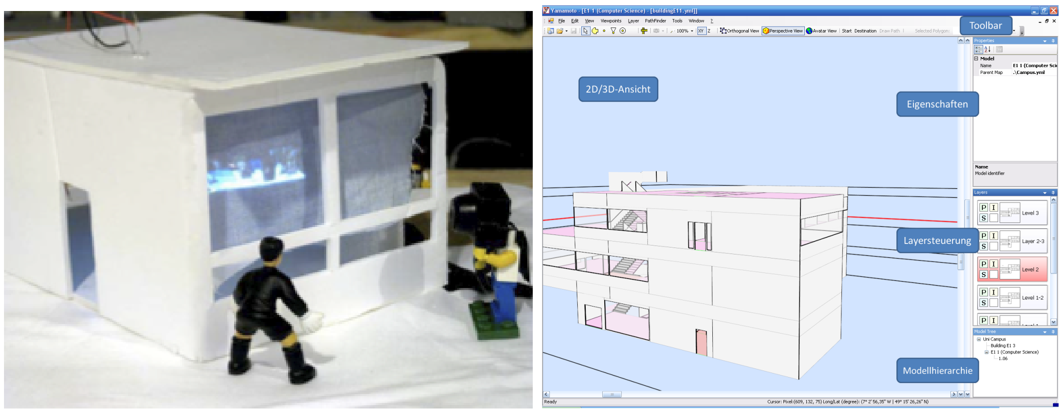


Figure 2.23: Left: A hybrid miniature model prototyping a media facade application using the City Compiler [134] Right: The Yamamoto editor for graphically modeling virtual environments [172, 174].

With the *Proximity Toolkit*, Marquardt et al. introduced a toolkit to simplify the exploration of interaction techniques by supplying fine-grained proxemic information between people, portable devices, large interactive surfaces, and other non-digital objects in a room-sized environment [117]. Their toolkit supports rapid prototyping of proxemic-aware systems and it includes different tools to observe, record and explore proxemic relationships within a 3D space. The toolkit is designed in a modulized manner and it separates sensing hardware from the data model. Hence, different sensing technologies can be substituted or combined to derive proxemic information. However, the main focus of the Proximity Toolkit is to model *proxemics* in an instrumented space. It is neither intended to universally prototype interaction in outdoor urban spaces, nor with architectural sized displays of different form factors and technical capabilities.

Stahl and Haupert introduced *Yamamoto* [172, 174], a map modeling toolkit that positions itself between 2D data structures — like street maps — or proprietary location models, and professional 3D CAD (Computer Aided Design) software for architects or photo-realistic visualization (see Figure 2.23). Yamamoto provides editor tools for graphically modeling virtual environments. Its main focus is modeling *location* within the scope of instrumented environments. Hence, it comes with a route-finding module called *Pathfinder*, which can calculate routes between any two points in the model. While Yamamoto could be used to model the space around a media facade, it is not capable of sufficiently modeling media facades installations with their unique characteristics, or interactions with them.

Halskov and Ebsen presented a conceptual framework for designing and prototyping complex media facades [83]. In the context of the use case of developing a 300m-long low-resolution curvilinear media facade, they address how content for a media facade needs to be designed with respect to the specific capabilities of the media facade and its interfaces. This includes scale, shape, pixel configuration, pixel shape and also light quality. However, they point out of that these properties can be addressed using a repertoire of existing design tools, tailored to specific purposes. They do not provide their own toolkit.

Texturing of 3D Objects

When developing virtual prototyping and simulation toolkits for interactive media facade installations creating a virtual 3D model of the media facade, its deployment environment, as well as the application deployed onto the media facade, the texturing of 3D objects is an important issue. First proposed by Catmull [39] in 1974, it has become a widely used and common technique. The term texturing denotes the process of covering the surface of a 3D object with 2D images, which are called *textures*. Heckbert defines a texture as *a detailed pattern that is repeated many times to tile the plane, or more generally, a multidimensional image that is mapped to a multidimensional space* [93]. He provides an overview on the fundamentals of texture mapping, including the geometric mapping, which wraps a texture onto an arbitrary surface. In general, for texturing 3D objects, we can follow two directions. We can either create a mapping of a 2D image onto the surface of the 3D object by distorting, scaling, rotating and moving the texture image or multiple copies of it until it covers the surface of the 3D object. Or, we can unwrap the surface of the 3D object onto a 2D plane which can be considered as a reverse mapping. Both approaches have in common that the textures are generally distorted when mapping them onto or unwrapping them from the surface of a 3D object. To address this issue, Maillot et al. introduce an approach to map textures with reduced distortion of the texture image. They construct an interactive texture tool for manipulating atlases in the texture space. In addition, they introduced an algorithm to automatically generate an atlas for various types of objects, in order to map different textures onto the objects. Stahl and Haupert illustrate a way to capture an application's visuals, mirroring it into a virtual display in real-time [173]. They argue that showing "the real outcome of the application logic" to a developer benefits both the design and testing process.

Visualizing Complex Content on Small Screens

When interacting with large contents like cartographic maps, web pages or 3D environments, the target area a person currently intends to interact with often covers only a small excerpt of the potential interactive content. Visualizing the whole content would therefore result in a waste of screen real estate. Common approaches to ease this problem are providing miniature representations or excerpts of the content. When dealing with geo-spatial content, a well known approach is to provide relevant excerpts of a map as orientation and navigation hints to the user. When interacting with 3D content or 3D environments, so-called world in miniature representations are often utilized. Stoakley et al. introduced the *World in Miniature* (WIM) metaphor [176]. They use a miniature copy of a virtual environment to create a second dynamic view-port onto a virtual environment in addition to the first-person perspective that is offered by a virtual reality system. The miniature copy is a scaled-down representation of the environment that can be manipulated in a single level of scale. This complicates the task of navigation through the WIM. In [199], Wingrave et al. addressed the problem of navigating and moving around in a WIM by adding scaling and scrolling to the WIM metaphor, which resulted in a *Scaled Scrolling World In Miniature* (SSWIM). However, since the WIM representation keeps the 3D form factor of a media facade in the miniature representation, such an approach is limited in its suitability for enabling continuous interaction with a media facade since — albeit virtually — the user still has to move around in the WIM in order to access different parts of the facade.

2.4 Summary

The work reviewed in this chapter reveals that so far, there is no common ground for designing and developing interactive installations for media facades in public settings. Existing installations are often carried out as design explorations with a playful or performative character. While existing installations often have a different focus, they turn out to suffer from the same challenges:

1. The public character of the setting has a strong impact on how people behave, how they interact with an installation and, due to the highly dynamic character of the setting, also on the general functionality of the installation as a whole.

2. The lack of development tools impedes the design process, since provided tools are usually tailored to one specific facade.

3. This results in interactive installations for permanent media facade being highly tailored to the particular media facade and usually not transferable to other media facades without further effort.

In order to exploit the full capabilities of media facades, we need a general understanding of the characteristics and properties influencing the outcome of an interactive installation for media facades. In order to analyze and evaluate interaction with media facades, we need to identify suitable measures to compare different settings and installations, as well as we need guidelines for the design of such. We need to develop new interaction techniques addressing the unique characteristics of media facades, as well as we have to adapt approved concepts from interacting with a digital display at-a-distance to the field of media facades. We further need general development and prototyping tools incorporating the media facade's deployment context and allowing to transfer applications between media facades.

3

Taxonomies of Interactive Media Facade Installations

Unlike designing the actual media facade — the underlying hardware — for a particular building, up to now, designing interactive media facade installations is far from being standardized. There is neither a common ground for which technologies and which interaction techniques work in a particular situation, nor is there any categorization that puts the components of media facade installations into context. Existing installations such as those described in Section 2.1 usually are one-of-a-kind deployments that (1) are specifically designed from scratch for the particular setting and (2) do not exploit the context of the setting. The design is usually driven by trying to make an existing design idea somehow fit into a given location, rather than analyzing the location and the context of the media facade to design an installation that specifically addresses the properties of the deployment environment. In this section we analyze media facade installations regarding the properties that frame the installation, their spatial context and the applied interaction techniques and utilized input devices. We further derive two taxonomies for describing interaction with media facades. The first taxonomy addresses the resolution of the media facade in relation to the resolution of the applied input technique. The second taxonomy for interactive media facade installations considers the technical and spatial properties of the media facade and its surroundings, as well as different input modalities. Our goal is to provide a set of measures and guidelines that inform the design of interactive installations for media facades in such a way that designers can derive a suitable interaction technique given a particular setting and design a particular media facade setting that is suitable for a particular interaction technique. Within this scope, we investigate different ways of participating in a media facade installation, and we analyze properties of media facades and their deployment environments in terms of their impact on interaction.

For the sake of simplicity, when addressing the people around a media facade, we denote them as *users*, without explicitly distinguishing between *passersby*, *potential users* and one particular *user*. Furthermore, with *media facade installation*, we refer to the notion defined as follows:

Definition 2. *The term **media facade installation** denotes the display of digital content on a media facade (as defined in Definition 1) within a public environment. A media facade installation comprises the displayed content, the underlying media facade, as well as the spatial setting around the media facade. A media facade installation is denoted as an **interactive media facade installation** if it allows passersby to interact with the displayed content.*

3.1 Taxonomy of Media Facades

Media facades are usually built to fulfill a particular purpose. They are based on different technologies and they are situated in unique, dynamic settings. This makes a comparison of different media facade installations a challenging task. As described in Chapter 2, there exist different approaches in the literature for categorizing media facades, resulting from different perspectives on the matter. A common perspective is to focus on technical, social and creative challenges of media facades. However, the absence of a clear boundary between media facades, media architecture and spatial art installations allows for even more different perspectives. As a first step towards providing a taxonomy of interaction with media facades, we categorize media facade installations based on the way in which they support user participation. We discuss the properties of a media facade installation that have an impact on the overall settings and derive a taxonomy for media facade installations, based on relevant properties of the overall setting. Haeusler categorizes media facades from an architectural perspective, based on the underlying technology. He distinguishes between *Mechanical Facades, Projection Facades, Window Raster Animation Facades* and *Illuminated Facades* [81] as follows (see Figure 3.1):

- *Mechanical Facades*

  Mechanical media facades physically alter the outer surface of a building. To create a mechanical media facade, the outer shell of a building — or part of it — is equipped with mechanical elements that can be kinetically changed to transform the facade of the building (as in Figure 3.1, (1)). Instead of displaying digital visual content, the main focus of mechanical media facades is displaying kinetic movements.

- *Projection Facades*

  For projection facades, the facade of the building remains physically untouched. It is digitally augmented by projecting content onto the facade. The projection can either be a front projection, where the content is projected from outside the building, or rear projection, with a projector inside the building, behind a translucent projection surface. Due to the use of projection, the resolution of the media facade depends on the resolution of the projector. Since projectors usually come with resolutions comparable to regular computer screens, the resolution of a projected facade is usually high enough to display fine-grained content (as in Figure 3.1, (2)).

- *Window Raster Animation Facades*

  For window raster animation facades, the individual windows of the underlying building are equipped with light emitting elements. Hence, every window is turned into a pixel of the overall media facade. The resolution of a media facade in this category therefore depends on the number of windows, which represents the number of pixels, as well as the size of the windows. As a result, this approach usually leads to media facades with resolutions of a few hundred pixels, limiting the complexity of content that can be displayed on the facade (as in Figure 3.1, (3)).

- *Illuminated Facades*

  The category of illuminated facades comprises the facades which are created by equipping the outer surface of a building with light emitting elements. Depending on the technology that is used, the media facade created can have various resolutions, ranging from low resolutions of less than a few hundred pixels to high resolutions comparable to common digital computer screens (see Figure 3.1, (4)).

Since the main focus of mechanical media facades is usually on displaying kinetic movements and not on displaying digital media content, and since kinetic architecture emerged as its own field of research, this thesis mainly addresses the categories of projection fa-

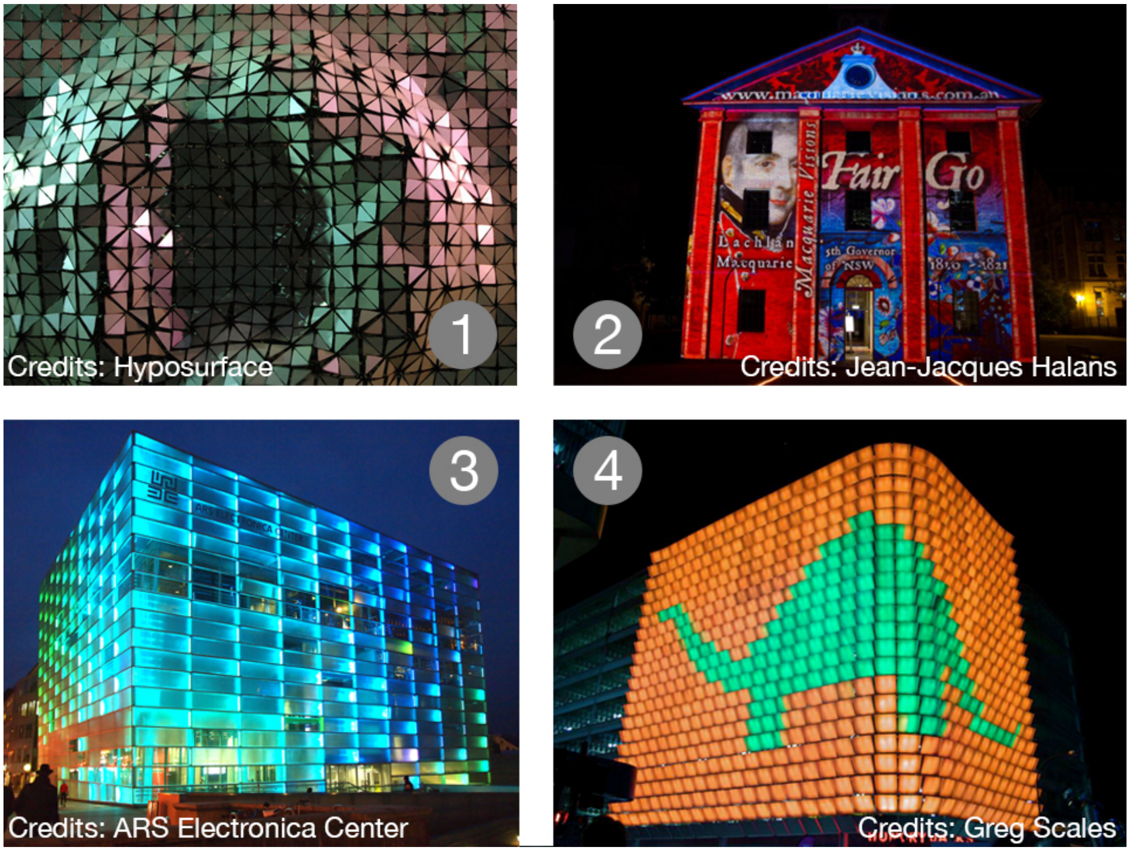


Figure 3.1: Four categories of media facades: (1) Mechanical facade, (2) projection facade, (3) window raster animation facade and (4) illuminated facade.

cades, window raster animation facades and illuminated facades. As we focus on interaction with media facades in this thesis, in addition to classifying media facades based on their technical capabilities, we need to put interactivity into context as well. With respect to this, we need to identify the parameters which determine the characteristics of interaction, as well as the technical capabilities and requirements involved in providing interactive installations for media facades.

Taking a more application-oriented approach, the Media Architecture Institute Vienna[1] categorizes media facade installations based on the purpose of the installation [182]. Their categorization does not consider particular forms of content in general. The only distinc-

[1] http://www.mediaarchitecture.org/

tion is made between static and dynamic content. In particular, they distinguish the following five categories of media facade installations:

1. **Animated Architecture**: Animated Architecture denotes media facade installations as well as media architecture, where animated elements are embedded into the architecture of a building without considering the underlying technology with which the installation is created. This considers both animation by mechanically moving elements and displaying animated digital content. Animated architecture hence includes light emitting media facades, as well as mechanical media facades.

2. **Spatial Media Art**: Media installations having a *spatial* component are referred to as Spatial Media Art. This includes voxel facades and digital light and media installations for spaces and places rather than the surface of a building. Besides the underlying technology of the installation, Spatial Media Art does not distinguish between static and dynamic content.

3. **Business and Money Architecture**: The category of business and money architecture is concerned with media installations addressing business and economic factors. This includes installations to branded buildings (i.e., hotels, corporate buildings or shopping centers) as well as to display advertisements, and basically all installations serving commercial interests.

4. **Participatory Architecture** considers installations allowing for participation in various ways, where the installation can react to changes in the environment and the space around the installation or also to individual persons. This includes both installations that react to the presence or movements of people in front of the installation, and installations providing means for explicit interaction between passersby and installation.

5. **Future Trends and Prototypes** finally summarizes installations such as kinetic facades, as well as prototypical installations which are not covered by any of the aforementioned categories. This category covers the wide spectrum of large-scale light installations and mechanical facades, but also small-scale installations and prototypes of novel technologies for media architecture in general.

	Interactive	Reactive (User)	Reactive (Environment)
Projected	Climate Wall [60], Night Lights [103], Spread.gun [56], SMSlingshot [55], TXTualHealing [103], Laser Tag [103], Wall of Light [103], Graffitti Dance [165], MobiSpray [157]	-	-
Illuminated	Aarhus by Light [30], Lummoblocks [80]	12m4s [103], La Vitrine Culturelle [82]	-
Window Raster Animation	Dexia Tower Touch [103], Blinkenlights [81], Marnix [82], Rundle Lantern [82]	-	Dexia Tower Weather Forecast [103]

Table 3.1: A basic media facade taxonomy with respect to the way in which a media facade installation allows passersby to participate. The taxonomy considers projected media facades, illuminated media facades, and window raster animation facades.

Since we are specifically addressing *interaction* with media facades in this work, we focus on the participatory aspects of media architecture to categorize media facades. Participatory media facade installations can vary in the supported types of participation. For *narrative* installations, the communication between user and media facade usually is an implicit, one-way communication from the installation to the user. The media facade communicates some information that passersby could consume. There is no limitation on the number of people that can consume the information and there is no explicit interaction between the media facade and a user. In this case, actions of users do not trigger any reaction of the facade. In contrast, participatory media facade installations can also support interaction in both directions between user and media facade. This includes *explicit interaction*, where users explicitly interact with a media facade (e.g., similar to interaction with a situated public display). More implicit ways of interaction are realized with *reactive installations*. In this case, the media facade reacts to passersby (e.g., the number of passersby, their moving direction, etc.) while they are not explicitly interacting with the media facade. In addition to reacting to passersby, a reactive media facade can also react to changes in its environment (e.g., weather conditions, traffic, etc). To address this, we further sub-divide reactive media facades accordingly into those that are reactive to particular users, or reactive to the environment around the facade as a whole. Table 3.1 depicts a basic taxonomy for participatory media facade installations based on the supported way of participation, as well as the technical realization of the facade. We categorize the media facade installations introduced in Chapter 2 with respect to the taxonomy. Since there exist multiple installations for permanent media facades, media facades can appear in more than one category. The taxonomy depicted in Table 3.1 can serve as a basic categorization of media facades based on the way in which they support participation. This distinction between reactive and interactive media facade installations is a very basic categorization. It does not further inform the actual scenario or the applicability of certain interaction techniques in particular settings. However, it still is an important aspect that we have to incorporate in further investigations. As we can see in Table 3.1, interactive media installations can be created with projected media facades, illuminated media facades and likewise window raster animation facades. All three categories of media facades constitute digital public screens at an architectural scale. The central differences characterizing the possibilities of interaction for interactive media facade installations in this case are the screen's resolution, as well as its pixel size, shape and density. One key characteristic of regular digital displays is that they are *well-defined*, which means they have equally-sized pixels which are distributed equally-spaced, resulting in a constant density of equally-sized pixels for each part of the screen. Since within the scope of this thesis, we investigate interactive media facade installations with media facades as digital public screens, we categorize media facades with respect to their potential to realize interactive media facade installations, as well as their ability to serve as a large-scale digital screen. Therefore, we define the category of media facades relevant for our work as *well-defined media facades*, which we define as follows:

Definition 3. *With the term* ***well-defined media facades****, we refer to projected media facades created by projecting content onto the surface of a building, window raster animation media facades created by turning the windows of a building into pixels and illuminated media facades created by embedding light emitting elements serving as pixels into the architectural structure of the building. Well-defined media facades are characterized by uniformly shaped pixels of equal size, which are arranged in a uniform, well-defined structure.*

3.2 Taxonomy of Interaction with Media Facades

While the taxonomy described in Section 3.1 categorizes media facade installations based on the underlying media facade itself, and the extent of interaction it allows for, the explicit interaction with media facades provides various characteristics to create a more fine-grained taxonomy for interactive media facade installations. Interactive installations for media facades usually apply different input modalities utilizing a wide range of input devices with various technical requirements. Furthermore, the technical properties and spatial arrangement of the interactive space itself influence the applicability of technologies to provide user input. In this section, we analyze and discuss the limiting factors for providing different forms of interactivity with media facades.

3.2.1 Resolution

Since media facades usually come with a wide range of resolutions, ranging from low-resolution facades, like the facade created for the Blinkenlights [81] installation with 144 pixels, to high-resolution facades like the facade of the Academy of Fine Arts Saar[2] with 2,400,000 pixels, the resolution is an important factor when designing interactive media facade installations. When displaying content on a digital display, the maximum resolution of the content is bound by the resolution of the display itself. While for regular computer screens the resolution is usually high enough that we can perceive the content even from very short distances (i.e., less than $10cm$), for large displays like media facades, the minimal required viewing distance to perceive the content as such constrains the distances from which a user can view and interact with the displayed content. When interacting with a media facade, the resolution or granularity of the user input is also an important factor for the interaction itself, the overall experience and the usability of the interactive installation. It influences whether the interaction is perceived as consistent when interacting with content of different complexity. To follow, we analyze both the role of media facade resolution and input resolution to put them in relation to each other. Furthermore, we derive a taxonomy for interactive media facade installations and categorize the interactive and reactive media facade installations reviewed in Section 2.2.2 accordingly.

[2] http://www.hbksaar.de

Resolution of Digital Displays

Besides the actual physical size, the general criterion for a digital display is its *resolution*. In general, the resolution of a digital display is the number of distinct pixels that can be displayed in each dimension of the screen. For a screen with a 2D form factor, the resolution is usually quoted as *width* x *height*. If, for example, the resolution of a display is quoted as *800*x*600*, this means the display has 800 pixels horizontally and 600 pixels vertically. Formally, the resolution of a digital display is defined as:

$$res = \text{horizontal pixels} * \text{vertical pixels} \tag{3.1}$$

When displaying content of different resolutions on a display, the technical realization of the underlying display technology influences the visual quality of the displayed content in terms of sharpness of the displayed content. Digital screens such as Liquid Crystal Displays (LCD) are built with a fixed raster of pixel elements, whose resolution is referred to as the *native resolution* of the display or simply as the *display resolution*. Since it is fixed due to the technical implementation of the display, an optimal display quality can only be reached if the resolution of the displayed content matches the native resolution of the display. Content with a different resolution than the native resolution of the display needs to be interpolated (e.g., scaled) to match the display's native resolution. This results in a loss of quality since the displayed content appears less sharp and visual artifacts might be introduced. In contrast, projected displays do not come with a fixed raster of pixel elements and can display content at different resolutions without interpolation. Besides the pure number of pixels a digital display has, the density with which the pixels are arranged is another common measure for the resolution. The *pixel density* sets the number of pixels of a display in relation to the display's physical size. Usually, the pixel density is measured in *pixels per inch* (*ppi*) and denotes the size of an image that can be displayed within a specified space. For example, an image of 100x100 pixels that is displayed in a 1 *inch* square has a resolution of 100*ppi*. The pixel density of a digital display can be calculated as follows: A common way to advertise the size of display is to name the size of the screen diagonal. Knowing the width w and height h of the display in pixels and the screen diagonal d_S in *inch*, the pixel density pd is given by

$$pd_D = \frac{\sqrt{w^2 + h^2}}{d_S} \tag{3.2}$$

Media Facade Resolution

The resolution of a media facade determines the appearance of the visual content displayed on it. Figure 3.2 depicts media facades of different resolutions, ranging from low-resolution facades with a few hundred pixels to high-resolution facades with millions of pixels. The complexity and the level of detail of the displayed content is also directly connected to the media facade's resolution. While low-resolution media facades usually display low-level, coarse-grained content like ambient visualization or animations, high-resolution media facades can display complex content, comparable to regular computer screens or situated public displays. In analogy to regular digital displays, we have to distinguish between the resolution of the media facade and the resolution of the displayed content. We define the terms *internal media facade resolution* and *external media facade resolution* as follows:

Definition 4. *With the term* ***internal media facade resolution****, we denote the resolution of the underlying pixel raster representing the media facade.*

Definition 5. *With the term* ***external media facade resolution****, we denote the resolution of the content that is displayed on the media facade.*

If a media facade is based on a fixed alignment of physical pixel elements, a mismatch in the internal and external media facade resolution results in interpolation of the displayed content. Since projected media facades are not produced by a fixed alignment of physical pixel elements, they can display content in different resolutions without interpolation. This results in a matching internal and external media facade resolutions. We refer to the internal media facade resolution with the term *media facade resolution.* For the sake of simplicity, if not stated otherwise, we assume that the internal media facade resolution equals the external media facade resolution and also that the content is displayed with the maximum physical resolution of the media facade and that it is not purposely downsampled to lower resolutions. As mentioned before, for digital displays with a planar, rectangular form factor, we can calculate the pixel density as a relation of the screen size and the number of available pixels (see Equation (3.2)). Since they are embedded into the architectural structures of a building and often cover more than one side of the underlying building, measuring the screen size of a media facade by its width and height or its display diagonal is not optimal. Instead, the surface area of a media facade can be used as a measure for its size. Hence, we define the pixel density of a media facade as a relation of the number of pixels covering the surface of the media facade and the size of the media facade's surface:

Definition 6. *The* ***pixel density*** *of a media facade is defined as the relation of the number of pixel elements n_p and the size of the media facade's surface A_M, measured in m^2. It can be calculated as:*

$$pd_F = \frac{n_p}{A_M} \tag{3.3}$$

Since there are no absolute measures for relating the pixel density of a media facade to the granularity of the displayed content, we can say as a general guideline:

Observation 1. *The higher the pixel density of a media facade, the higher the granularity of details that can be displayed. The lower the pixel density of a media facade, the lower the maximum complexity of content that can be displayed.*

Depending on a display's size, its shape, resolution and pixel density, a person needs to keep a certain minimal viewing distance to be able to comprehend the content as a whole [98, 178, 1]. At the scale of media facades, this becomes even more important since it influences the kind of content that can be displayed as well as its granularity and the level of detail. For ambient visualizations and pure animation, the required minimal viewing distance only plays a subordinate role. As the term *ambient* denotes, in this case the content — or a part of it — is supposed to be consumed passively. With an increasing resolution of the media facade and the resolution of the displayed content, the granularity and the level of detail of the communicated information — from the media facade to passersby — can increase. Accordingly, the required minimal viewing distance, and perceiving the content, becomes more important. Since the required minimal viewing distance depends on a variety of factors, we cannot calculate an exact value for the required minimal viewing distance for a given media facade installation. However, leaving the placement of the facade aside, we can state that:

Observation 2. *The higher the pixel density of a media facade, the lower the required minimal viewing distance from which two pixels can be distinguished by a user. The lower the pixel density of a media facade, the higher the required minimal viewing distance from which two pixels can be distinguished by a user.*

Figure 3.2: Media facades of different sizes, shapes and resolutions.

Input Resolution

When designing interaction with media facades, besides the resolution of the media facade itself, the resolution and granularity of the utilized input technology are important factors as they determine the precision of user input. As previously mentioned, within the scope of this thesis, we address continuous or ongoing interaction with a media facade. In this respect, we do not consider the simple provisioning of content (e.g., pushing information such as stock prices onto a facade). From a user perspective, while interacting with the content of a media facade, the resolution of input should at least match the resolution of cursor movements on the facade in order to create a consistent and fluent interaction experience and to avoid aliasing effects that could result in a stuttering movement of the cursor. A severe mismatch in the resolution of input and the resolution of the triggered cursor movement on the facade can cause a mismatch in the user's perception and hence increase the mental workload while interacting [38]. With 3D content, this becomes even more prominent [110]. However, a mismatch in the resolution of both user input and media facade can also be explicitly intended by a designer as a feature of the installation. E.g., a mismatch in the resolution of user input and media facade could be explicitly created to emphasize particular characteristics of the interaction. This is a popular approach when designing interactive installations encompassing performative or playful elements, usually utilizing whole body interaction. For example, when basic body movements such as pointing to the left or right are tracked to determine the movement direction of continuously moving objects displayed on a high-resolution media facade.

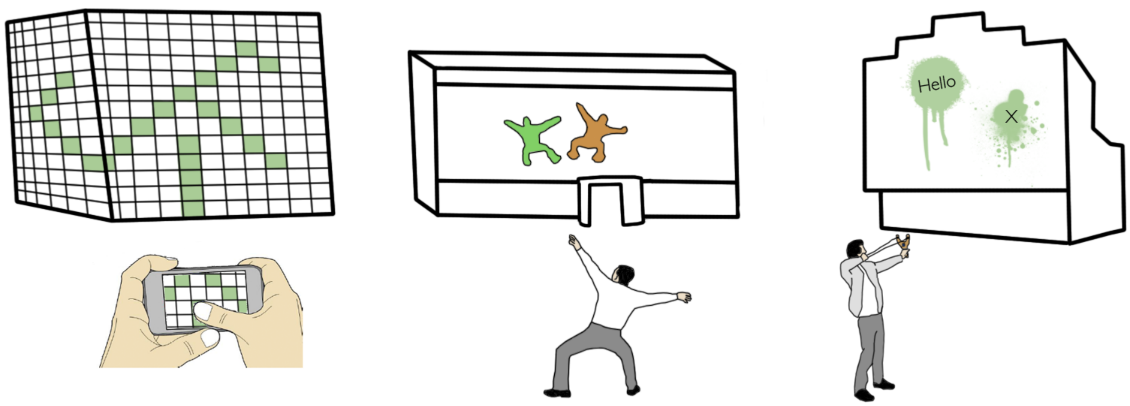


Figure 3.3: Illustration of the resolutions of media facades and different interaction techniques. Left: A low-resolution media facade with high-resolution touch input. Middle: A media facade with medium resolution supporting whole body interaction. Right: A high-resolution media facade supporting direct pointing.

Relating Input and Media Facade Resolution

Figure 3.3 illustrates the variance in resolution of media facades and different interaction techniques and input modalities. While for the media facade the resolution can be measured in the number of pixels relative to the actual size of the facade, measuring the resolution of user input is more complex. While the resolution of the input technique is given by the utilized input devices, the resolution of the actual user input depends on the perception of the particular user. As depicted in Figure 3.3, the interaction can involve different input devices, which usually come with different input resolutions. However, different input devices and techniques can also be combined. An example for such a combination is the SMSlingshot presented by Fischer et al. [55]. This installation combines direct pointing with body movement to allow for user input. A laser pointer is attached to a wooden slingshot with which the users have to perform an *aim-and-shoot* gesture to shoot colored text messages onto a projected high-resolution media facade. As previously mentioned, the input resolution perceived by a user while interacting with a media facade can differ from the actual resolution of the utilized input device. In the case where the interaction with the media facade is not bound to a fixed position within the display space, an additional factor is introduced. If the interaction can be carried out in a way that allows the users to walk around the potential interaction space while interacting, the perceived resolution of the input and of the media facade can change. For direct interaction techniques such as interacting by pointing at the media facade, with an increasing distance between user and media facade, the sensitivity of the input — and hence the perceived input resolution — increases. The movement angle of the pointing device results in larger pointer movements on the screen. In general, we can assume that:

Observation 3. *With an increasing distance between a user and the media facade, the perceived resolution of direct input techniques (e.g., direct pointing at the facade) will increase.*

In the Urban HCI model [54], Fischer and Hornecker describe the potential interaction space of interactive media facade installations as the space from which interaction with the media facade can be potentially carried out. Considering the dynamic change of the perceived input resolution when moving around within the potential interaction space while interacting with the media facade, the perceived input resolution should also be a limiting factor for the size of the potential interaction space. Interaction with the media facade should be perceived similarly from any position within the potential interaction space. However, since the perceived input resolution from different positions and distances depends on the subjective impression of the particular user, we cannot quantify the perceived input resolution with absolute metrics for defining a threshold for a maximum distance within the potential interaction space from which interacting with the media facade should be allowed for having a consistent interaction experience. Independent from the perceived input resolution while interacting, for interactive media facade installations in general we can have the following three scenarios:

1. The resolution of the input and of the media facade match,
2. the input resolution can be more fine grained than the facade's resolution or
3. the facade's resolution can be higher than the input resolution.

We can express the relationship between the resolutions of the input and the media facade with a *resolution factor Res*, which we can determine as follows:

$$Res = \frac{Input\ resolution}{Media\ facade\ resolution} \tag{3.4}$$

where *input resolution* is defined as:

Definition 7. *The input resolution denotes the granularity of user input. The input resolution is composed of the resolution of the input device and the accuracy with which a user operates it.*

To allow for fine-grained interaction with a media facade in relation to its resolution, we need a resolution factor of $Res \geq 1$ (as in Figure 3.5, green area). In this case, continuous input such as controlling a pointer using a pointing device can be smoothly mapped to the respective pointer movement on the media facade without jumping. While increasing the accuracy of the utilized interaction technique increases the possible maximum accuracy of input, an increasing resolution of the media facade increases potential complexity and the level of detail of the displayed content. As previously mentioned, we have to distinguish between the accuracy of input as a matter of the user's subjective perception and the actual accuracy of the input technique or the input device. The resolution of the interaction technique — or depending on the actual technique, the resolution of the input device — represents an upper bound for the actual input resolution. The accuracy of input can be perceived differently, depending on the subjective impression of a user and also on the accuracy with which he or she operates the input device (see Definition 7). In the optimal or at least most natural case, the resolutions of input and media facade result in a resolution factor $Res = 1$ [38]. In this case we can provide a fluent and consistent interaction, where user input is correctly mapped to actions on the facade. In this respect, when increasing the resolution of the media facade or the complexity of the displayed content, we should also increase the resolution of the input to provide a consistent experience. To illustrate this, we can consider the following example: A user is painting onto a media facade using the touch screen of a mobile device as the input device. As depicted in Figure 3.4, we consider two different ways of painting by controlling a pointer on the facade: interaction through live video (i.e., direct interaction with a spatial mapping) and using keys on the mobile device (e.g., up, down, left, right) to control a pointer on the facade (i.e. indirect interaction without an explicit spatial mapping). In the case where there exists a spatial mapping between input (e.g., the user's input on the mobile device) and the output of the facade (see Figure 3.4), to provide a consistent interaction experience, for an increasing resolution of the media facade we should also increase the resolution of the input device as mentioned previously. If there is no spatial mapping between input and output as depicted in Figure 3.4 (left), an increasing difference of input resolution and media facade resolution comes at the cost of sequencialization of input. This means that in order to perform the same movement as in the case of a spatial mapping (see Figure 3.4, right), the movement is split into a sequence of movements. The length of this sequence — the number of steps — depends on the difference in resolution of input and media facade. With an increasing difference in input and facade resolution, the length of the sequence increases. This can be counterbalanced by respectively increasing the resolution of the input device.

We can consider a resolution factor of $R = 1$ as a lower bound when aiming for comparably fine-grained interaction with complex content. In the case where the resolution is higher than the resolution of the media facade, the media facade's resolution limits the input resolution perceived when interacting with the facade. However, a higher resolution of the input device can also be exploited to augment the interaction. For example, when using the high-resolution touch screen of a mobile device as the input device for a low-resolution media facade, we could provide additional local controls or local content with a higher level of detail. If the resolution of the media facade is higher than the resolution of the input ($Res \leq 1$), we cannot allow for a comparably fine-grained interaction with the complexity of the displayed content given by the resolution of the media facade. However, if the intended interaction has a playful or performative character, or contains artistic elements, this could be also sufficient to meet the requirements of the interaction.

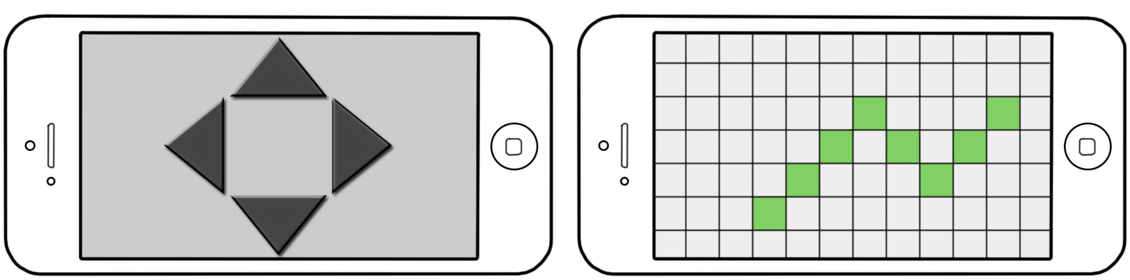


Figure 3.4: Using a mobile device as an input device for controlling a pointer on a remote display with (right) and without (left) a spatial mapping between input and output.

If we revisit the interactive media facade installations reviewed in Section 2.2.2 and compare the resolution of the media facade with the resolution and granularity of the user input, we obtain a taxonomy as depicted in Figure 3.5. The diagram does not use absolute scales. The goal of this taxonomy is to relate input resolution and media facade resolution to each other. As can be seen in Figure 3.5, aligning the interactive media facade installations reviewed in Section 2.2.2, the taxonomy reveals three clusters:

(A) **Fine-grained interaction for high-resolution facades:** It turns out that for interactive installations allowing for fine-grained interaction with complex content, the input resolution indeed is usually at least as high as the media facade resolution, with $Res \geq 1$ (see Figure 3.5 (A)).

(B) **Fine-grained interaction for low-resolution facades:** For interactive installations for media facades with low resolution, the resolution of input and the granularity of the interactions is usually approximately the same as the media facade's resolution, with $Res = 1$ (see Figure 3.5 (B)). Although there exist installations such as *Dexia Tower Touch* [81], providing high-resolution input devices (e.g., a touch terminal) for a low-resolution facade, the granularity of the actions triggered on the facade is still limited by the facade's resolution.

(C) **Coarse-grained interaction:** For performative interaction and interaction with a playful character, we observe a mismatch in the resolution of the input and the media facade with $Res \leq 1$ (see Figure 3.5 (C)). As pointed out before, this is usually intended as a part of the performative or playful interaction.

3.2.2 Environmental Context

As previously mentioned, media facades are usually situated in a dynamic setting with unique characteristics making it difficult to compare installations and situations that arise from them. Nevertheless, there are characteristic properties in the environmental context of a media facade's deployment space which have a strong influence on the applicability of different technologies. These properties can be brought in to categorize and compare different media facade installations in different contexts.

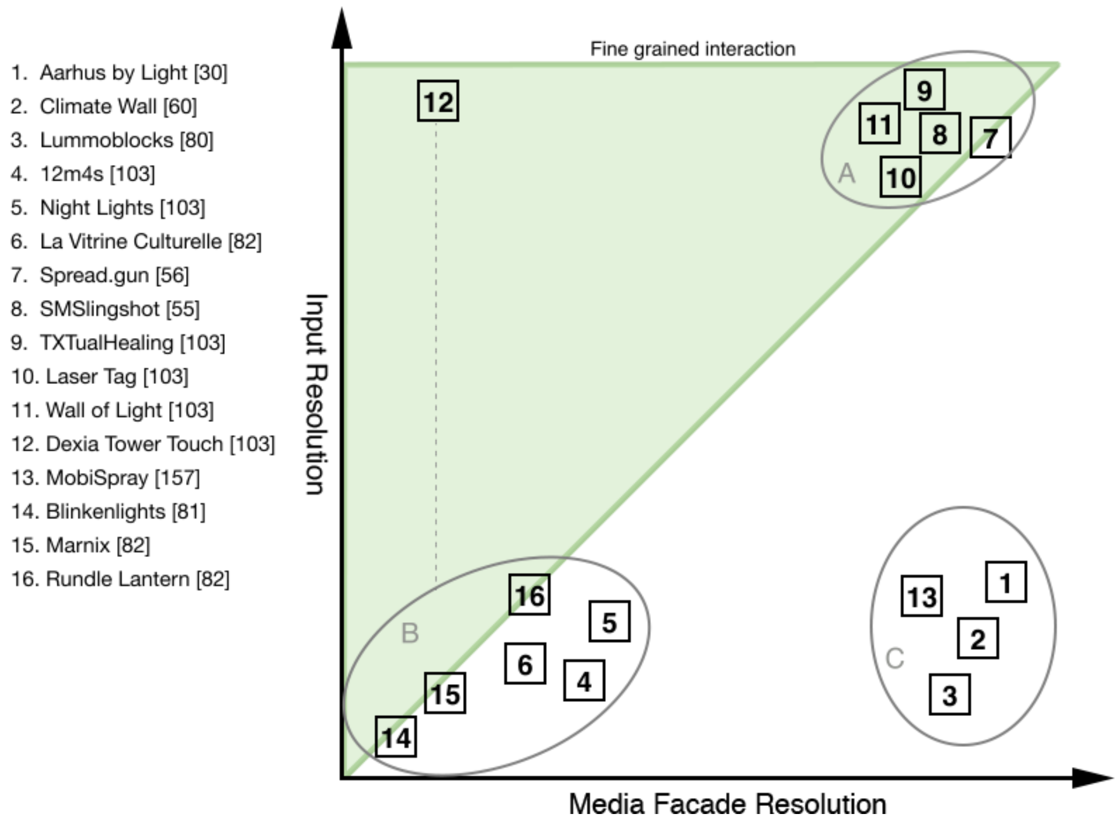


Figure 3.5: The taxonomy of interactive media facade installations based on the resolution of user input and the resolution of the media facade itself. We categorized the interactive installations reviewed in Section 2.2.2 accordingly. (A) The input resolution matches the facade resolution and both have high resolution. (B) The input resolution matches the facade resolution and both have low resolution. (C) The facade resolution is higher than the input resolution. Green area: The area with a resolution factor of $Res \geq 1$.

Spatial Aspects

As described in Section 2.1.1, Fischer et al. analyzed spatial aspects in the design of shared encounters for interactive media facades with respect to the spatial configuration in relation to the structuring of interaction [54]. In combination with the resolution of the media facade itself, as well as its underlying technology (e.g., pixel-based, projected, etc.), the configuration of the space around a media facade has more impact on the appropriateness of content than on the general applicability of particular interaction techniques. Size and configuration of the space determine distances — including minimum and maximum viewing distance — and angles from which users can see the facade and the displayed content. To allow the users to perceive the displayed content as a whole from within this space, the content must be designed with respect to the spatial configuration of the setting. When dealing with multiple users interacting simultaneously with the facade, the spatial configuration of the potentially interactive space can further determine if the particular users can see each other directly or if they are at least aware of each other. The spatial aspects also influence the communication between the users and hence, the social configuration of the setting.

Concerning different interaction techniques, the spatial configuration of the setting represents a scaling factor rather than a criterion for the general applicability of a particular technique. From a spatial point of view, interaction techniques in the context of media facades in general only require minimal space — the immediate space around the user — to be applicable at all. For interaction that does not involve whole-body movement of the user, the minimal space is usually represented by the space required to operate the input device. When involving interaction techniques such as direct pointing, it further requires a direct line of sight between the user and the particular target area on the facade. Concerning whole body interaction, size and spatial arrangement of the involved spaces determines the possible scale of the mapping between body movement and reaction on the facade. Furthermore, it influences the potential number of maximum users that can interact simultaneously, since each interacting user requires a certain amount of the available space. In summary, we can say that the spatial configuration has no direct influence on the general applicability of a particular interaction technique. Hence, we only assign it a minor role in our taxonomy.

Instrumentation

The technical instrumentation of a media facade's environment has a direct influence on the applicability of different interaction techniques. Different interaction devices and interaction styles require the instrumentation of the potential interaction space to different extents. To follow, we point out the influence of different characteristic properties of interactive media facade installations on the technical instrumentation of the potential interaction space.

Interaction Style: Interaction with media facades can be designed in various ways, requiring the potentially interactive environment to be instrumented to a different extent:

1. *Performative interaction* is usually realized by tracking the body posture or movement of a user. This is usually achieved by instrumenting the potentially interactive space with additional hardware such as cameras or sensors.

2. *Direct Interaction* — such as direct pointing — requires either tracking the current position of the user, as well as the direction he is pointing at, or tracking the location of a cursor on the facade, initiated by a pointing device (e.g., a laser pointer). The latter is usually realized with cameras pointed at the facade and hence requires less instrumentation than tracking a user's body movements within a large open space.

3. *Indirect* or *relative interaction* techniques (e.g., using a mobile device's built-in inertial sensor to relatively control a pointer on the facade) usually do not require an extensive instrumentation of the potentially interactive space. The technical capabilities required to determine the user's input are usually provided by the input device itself (i.e., a mobile device or an interactive terminal placed in the space in front of the facade).

Connection: When using dedicated input devices (e.g., custom-built hardware or smartphones), there is the requirement for a permanent data connection between input device and media facade. With the increasing availability of public *WiFi* networks and permanent data plans for mobile devices, the demand for connectivity can be considered as a minor concern, and not as a dedicated instrumentation of the space around the media facade.

Number of Users: Aside from the interaction technique itself, the number of supported users is influenced by the technical instrumentation of the potentially interactive space. The instrumentation can be either a permanent installation of static hardware, the distribution of mobile input devices, or both. For example, if we consider whole body interaction where the user's movements are tracked with a camera, the number of users is limited to the number of users trackable with the available cameras. To allow for more users, the number of cameras has to be increased. For the sake of simplicity, we omitted the required computing power to process the actual camera data in this example, since this again depends on the particular algorithms, update rates and the detailed realization. Similarly, when using dedicated input devices, the number of users is limited by the number of provided devices and can be increased by distributing more devices up to a certain limit given by the application itself.

Interaction Distance: As mentioned before, interacting with a media facade usually implies interacting at a distance. As can be seen when comparing techniques for interacting at a distance, such as *SMSLingshot* [55] or *LaserTag* [103], they require different minimal viewing distances and they can potentially work from great distances. Depending on the desired interaction distance, this can require the instrumentation of large spaces. On the flip side, if a designer aims for interaction from very short distances within the direct proximity of the facade, some interaction techniques are not applicable if the intended interaction distance is below the required minimal distance.

Location Awareness: The degree to which an installation can be aware of a user's position within the potential interactive space is also determined by the instrumentation of the environment. In the least constrained case, the application only needs to be aware of whether a user is within the potential interaction space or not. Since this is covered anyway by the chosen interaction technique, as a user needs to be within the potential interaction space for the technique to work, we can treat this as though the installation not being aware of the user's location and this case therefore does not need further instrumentation. If the facade needs to be aware of whether a user is within its direct proximity — which is often the case for interactive media facades installed at ground level — only the immediate space around the facade needs to be instrumented. If the exact location of a user within the potential interaction space needs to be known, the amount of required instrumentation increases with the size of the space and of course the accuracy with which the location needs to be provided.

The spatial configuration of the potentially interactive space again only plays a minor role for the general possibility of instrumenting it with different technologies. The size of the space simply indicates the amount of resources required. Or, conversely, the available amount of resources determines the size of the space that can be instrumented.

3.2.3 Input Modalities

When designing interactive installations for media facades, the applied interaction techniques in combination with the utilized input devices shape the interaction and have a huge impact on the usability and the user experience. Furthermore, they determine the technical requirements for the instrumentation of the potential interaction space in front of the facade. When choosing a particular input modality (e.g., touch, pointing, etc.) and a particular interaction technique (e.g., camera tracking, interaction through live video, etc.) there is always a trade-off between the technical requirements for utilizing the particular input modality and the granularity of the input and the triggered experiences. Hence, the utilized input modalities are an important factor for a systematic approach to categorize interaction with media facades.

Whole Body Interaction

In Section 2.2.2, we reviewed interactive media facade installations using whole body interaction to allow passersby to control content on a media facade. While the human body itself is used as an input device, such an approach requires tracking body movement within the potential interaction spaces. Hence, those spaces have to be instrumented with different technologies, depending on the accuracy and the extent to which movements are supposed to be tracked. With the instrumentation of the space, and turning every body entering the potential interaction space into an input device, this input modality is especially suited for reactive media facade installations where the communication with the facade happens implicitly. Whole body interaction also turned out to be well suited for performative interaction within an artistic or playful context [30, 60, 81, 82]. In this context, mapping body movement within a particular space — physically running around within the space — to motion on the media facade is a common pattern. This allows for a continuous interaction through the whole potential interaction space. Furthermore, to passersby, users seem to *perform* in front of an audience, which usually draws attention and encourages passersby to engage in the social and physical setting [54, 30].

For whole body interaction, the accuracy of the tracking and hence of the interaction itself heavily depends on the utilized tracking technologies, the speed of the movement, and environmental conditions such as lighting conditions. From this, we argue that whole body interaction is more suited for coarsely grained interaction where the accuracy of input is not of great importance. This is especially the case for low-resolution media facades that are not capable of displaying complex content. For fine-grained, precise interaction with complex content, we need more accurate interaction modalities. Concerning the number of users, whole body interaction is suitable for multiple users interacting simultaneously. It is only limited by the technical capabilities of the utilized tracking technologies and the size of the instrumented space.

Pointing

Direct and indirect pointing are further common techniques for interacting with a media facade (see Section 2.2.2). Both usually require a dedicated input device with which the user points at the facade in order to interact (e.g., to control a pointer). While for direct pointing the movement of a visual cue that is initiated by the pointing device — usually a laser pointer — is tracked on the facade with the help of cameras, in the indirect approach, the input device itself calculates the relative movement — or in some cases the position — of the pointer and communicates it to the facade. Hence, the amount of required instrumentation of the potential interaction space is rather low, reducing the effort needed to set up the installation, while allowing for fine-grained interaction with complex content at the same time. The low instrumentation makes such an approach highly suitable for portable, non-permanent media facade settings. It is common to use direct or indirect pointing in combination with non-permanent, projected media facades. In the case of direct pointing, a camera can be mounted onto the projector for observing the facade and tracking pointers. In the case of indirect pointing, the pointing can be realized with a portable mobile input device while no further hardware — the projection system aside — is required. It is common to use a projector which is connected to a mobile computer. For direct pointing, the most common pointing device is a laser pointer which also might be mounted onto or integrated into custom input devices. The position of the pointer on the projected facade is then usually determined with the help of a camera mounted onto the projector. For indirect pointing, a common input device is a standard mobile phone, which allows for calculating the relative pointer movement based on the movement of the device itself or based on touch input on the mobile device's screen. Since in both cases dedicated input devices — usually portable — are needed, such an approach usually needs an ongoing monitoring of the setting. Although multiple users could interact with the media facade in parallel, this is hampered by two factors: (1) Every user needs an input device. Hence, the number of potential users is limited by the number of input devices. However, if the user's own mobile phone is used as the input device, then this is less of an issue. (2) Furthermore, when using direct pointing with a laser pointer-like device, distinguishing between the respective users is challenging. Such a system usually only detects a certain number of laser pointers on the screen, but it is not aware of how many lasers they are coming from and which pointer belongs to which user. To match a pointer to a particular user, further instrumentation of the input device is needed.

Touch and Gestures

Gestural user input by direct touch is a widespread approach often used with permanent installations (see Section 2.2.2). When applying touch as the input modality of choice, dedicated input devices are needed. This is a common approach utilized in various permanent media facade installations of various resolutions [81, 82]. For such installations, a stationary input device (e.g., a input terminal or an interaction booth) is usually provided within viewing distance of the facade. While this comes with the benefit of keeping the required instrumentation at a minimum (e.g., only the stationary input device needs to be installed and usually no instrumentation of the further environment is required), this comes with the disadvantage of usually limiting the number of users to one or two. Furthermore, this leads to stationary settings, not exploiting the spatial aspects of interaction with media facades. While interacting, users are required to be within the direct proximity of the stationary input device. This can also turn into an advantage when realizing complex interactions requiring specialized or dedicated input hardware to address the peculiarities of a particular setting. Concerning the resolution of the facade, this approach is in general highly suitable for both low- and high-resolution facades.

Web and Remote Interface

A further approach to allow users to interact with the content of media facade is providing external — usually web-based — interfaces to upload and modify the displayed content. While such an approach does not require instrumenting the environment in front of the media facade, it also does not particularly address the passersby in front of the facade within the display space. Furthermore, users do not get immediate visual feedback on their interactions. The main focus of such approaches lies in supporting a one-directional interaction, allowing a wide range of users to provide content for the media facade pushing this approach more in the direction of being reactive rather than being interactive. By providing remote interfaces, the target audience for interacting with the facade facade can be maximized since basically every person around the globe with access to the internet could potentially interact. However, we believe that this does not go along with the main idea of interacting with a media facade. It rather constitutes pushing content onto a media facade and we consider this as the minimal amount of participation that can be allowed.

3.2.4 Taxonomy

One goal of this thesis is to provide a taxonomy of interaction with media facades. The aforementioned properties and characteristics of an interactive media facade installation determine the applicability of different approaches for allowing interactivity. Furthermore, we can consult the taxonomy to compare interactive media facade installations and also to design an installation for a given environment. In general, we address the following properties:

- **No instrumentation, versus instrumentation** of the potential interaction space: The technical realization of interaction can require instrumentation of the potential interaction space in front of the facade to different extents. This can vary from providing custom input devices to equipping the space with complex infrastructure.

- **Utilized input modalities**: Different input modalities usually come with different technical requirements. They vary in the level of user engagement they require. Furthermore, depending on their characteristics, there are input techniques more suitable for high-resolution media facades and some that work well with low-resolution facades.

- **Stationary versus mobile** interaction: Interaction can be carried out in either a stationary or a mobile way. For stationary interaction, users are either provided with a stationary input device, or the interaction requires the user to remain in a particular place. When mobile interaction is possible, users usually can move around within the potential interaction space and interact from arbitrary places. Furthermore, they often can continuously interact while moving around the potential interaction space. The particular movement of the user can also be a part of the interaction itself.

- **Resolution** of the media facade: Media facades can have different resolutions, ranging from very low resolutions of less than 200 pixels to high-resolution screens with several million pixels (see Section 3.2.1). This for example influences the options for displaying visual feedback such as pointers or controls necessary for particular input techniques, as well as the complexity of the displayable content. Furthermore, as described in Section 3.2.1, when the interaction is carried out while moving around within the potential interaction space, the perceived resolutions might change dynamically.

- **The media facade's form factor**: Different forms of interactivity can involve the presence of a user in various ways. Users can be required to be within a certain distance from the media facade in order to interact with its content in general. A user could also explicitly interact with a media facade by changing his distance from the facade. Furthermore, when providing means for remote interaction, a media facade might not require a user to be present at all. When considering the media facade as a whole and not only the particular part visible from a user's current point of view, the form factor of a media facade can introduce additional complexity. With interaction techniques requiring a direct line of sight between the user and the target area on the facade, interacting with all parts of the facade is not possible without relocating and moving around the facade. Furthermore, the media facade's content might never be completely visible as a whole to a user, which also can constrain the interaction.

The aforementioned properties are dependent on each other in different ways. We combine them into a taxonomy of interaction with media facades as follows: We initially classify the interactive installation according to the utilized input modality and whether it requires the instrumentation of the potential interaction space or not. Regarding input modalities, we consider whole body interaction, pointing, touch and gestures and remote interaction as described in Section 3.2.3. We subdivide the installations into stationary ones, where the position from which a user interacts is fixed, and mobile ones, where a user can interact from different locations and also especially when moving between location. For the input modalities, we subdivide the applicability of the particular input modality for low- and high-resolution media facades. For each installation, we point out the number of supported users. This results in the spatial layout of the taxonomy as depicted in Figure 3.6. The layout was inspired by the taxonomy for mobile device input on external displays, presented in [20].

3.2.5 Classification of Existing Approaches

Based on the introduced taxonomy, we revisited existing interactive media facade installations presented in Section 2.2.2 and categorized them accordingly as depicted in Figure 3.7. It turned out that whole body interaction usually demands static instrumentation of the potential interaction space, usually resulting in interaction from a static position with a limited number of users. Furthermore, the classification revealed that

Figure 3.6: The spatial layout of the taxonomy of interaction with media facades.

whole body interaction is usually applied in combination with low-resolution media facades. Pointing as well touch input also require the instrumentation of the potential interaction space. Even in the case of using a laser pointer as the pointing device, a camera is still required to track the pointer movement. While pointing installations usually address high-resolution media facades, touch and gestural input are regularly applied for both low- and high-resolution media facades, which we believe involves using additional input devices, handling the mapping between input and facade resolution. Manipulating a media facade's content through remote interfaces (e.g., web interfaces) does not require the instrumentation of the space in front of the media facade, but it usually also does not require a user to be present within the display space near the facade. The classification reveals that for a *non-instrumented* potential interaction space, no system exists that considers *pointing* at the facade or allowing direct *touch input* from *mobile* positions in front of the facade. One expected outcome of this thesis is to provide lightweight interaction techniques for these categories (see Figure 3.7, gray area), which do not require an instrumentation of the potential interaction space and allow for fine-grained interaction with media facades of arbitrary form factors and resolutions from mobile positions within the potential interaction space. We believe that this a key aspect for turning media facades into fully fledged interactive surfaces in urban environments.

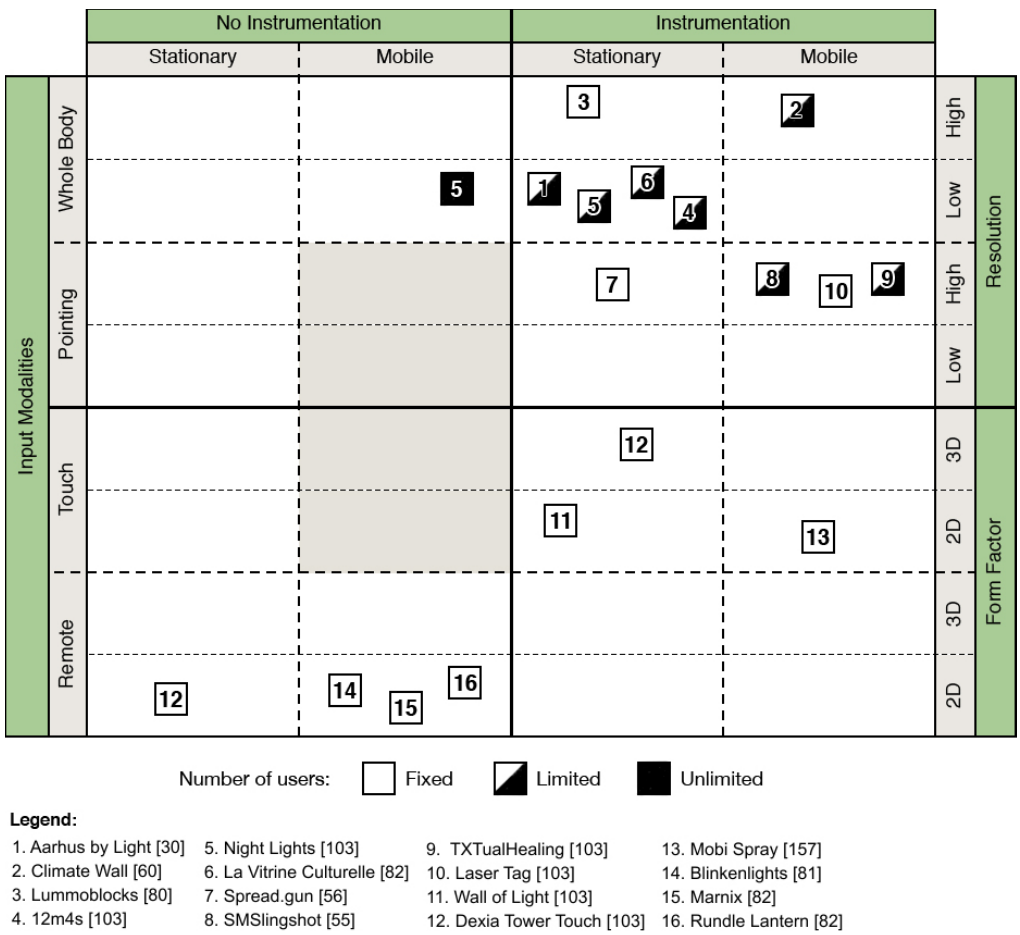


Figure 3.7: Classification of the interactive media facade installations reviewed in Section 2.2.2. Installations can appear in more than one category. The classification reveals that for a *non-instrumented* potential interaction space, no system exists that considers *pointing* at the facade or allows direct *touch input* from *mobile* positions for media facades with arbitrary form factors (indicated by the gray area).

3.3 Implications on the Spatial Setting

The spatial setting of an an interactive media facade installation is not only bound by the physical structure of the space, but also by the characteristics of the media facade itself, as well as the implemented interaction techniques. The size and resolution of the display, the applied interaction technique, as well as the space around the display likewise shape the particular spaces. The display size in combination with the interaction technique determines the distance from which a user can interact with the display within the potential interaction space. As a result, this further influences the appearance and size of possible gap spaces within the display space. E.g., for direct touch input there will be no gap space, but when utilizing mobile devices as input devices there will be a gap space depending on the way the mobile device is used to generate input for the display. The introduced gap spaces can either be introduced by the interaction technique itself or they can be physical, caused by obstacles within the physical environment around the display.

In the following example, we apply the Urban HCI model [54] to the same media facade for three different interaction techniques to illustrate the influence of the interaction technique's characteristics and requirements on the potential interaction spaces. The interaction techniques comprise a stationary interface, a mobile interface, and whole body interaction realized by camera tracking. For the example, we chose a media facade in the city center of Sao Paulo, Brazil (Location: Avenida Paulista, Sao Paulo, Brazil. GPS: -23.555547, -46.663016). As we can see in Figures 3.9, 3.8, 3.10, depending on the applied interaction technique, the spatial configuration of an interactive media facade installation can change drastically. In this setting, the display space includes a large number of gap spaces that are introduced mostly by streets crossing the space, causing the spaces that could serve as potential interaction spaces to be limited in size and distributed across the display space. For better readability, we did not mark the streets explicitly as gap spaces in Figures 3.9, 3.8, 3.10. As an additional difficulty of this setting, the non-gap spaces intersect with the regular passageways, with passersby crossing the setting. For a mobile interface (compare Figure 3.8) with which users can move around, we can see that non-coherent potential interaction spaces can arise, which are distributed within the display space. As a result, the potential interaction space and hence the potential number of users could be maximized.

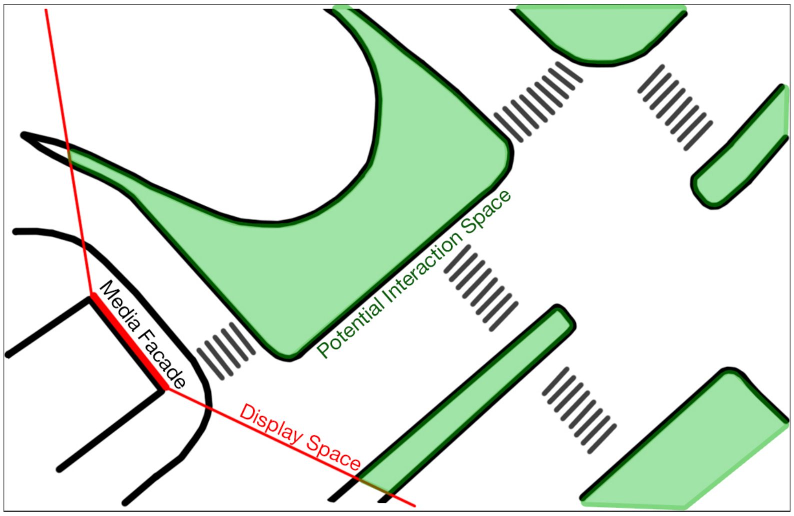


Figure 3.8: The spatial configuration for a movable interface according to Fischer and Hornecker [54]. For a better readability, the streets are not marked as gap spaces.

For a stationary interface (see Figure 3.9), the places within the display space that could potentially host the interface and with this form the potential interaction space are few. Since the interface is installed permanently, in addition to the gap-spaces, the spaces covering the regular passageways are also excluded to avoid blocking the stream of pedestrians.

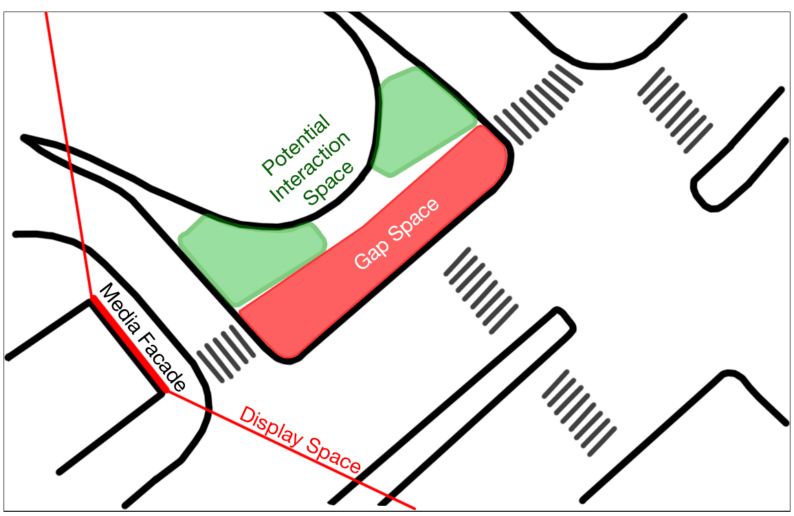


Figure 3.9: The spatial configuration for a stationary interface according to Fischer and Hornecker [54]. For a better readability, the streets are not marked as gap spaces.

For a camera interface (see Figure 3.10) that tracks a particular space to allow whole body interaction similar to the Lummoblock [80] installation, arrangement of the potential interactive spaces further depends on the implemented interaction style (see Section 2.2.2). If the interaction does not require a user to physically move around to interact, the spaces are similarly shaped as in the aforementioned case of mobile interfaces (see Figure 3.8). If the interaction requires a user to physically move around to interact, the spaces are similarly shaped as in the aforementioned case of stationary interfaces (see Figure 3.9), with the amount of movement as an additional constraint. For example, Lummoblocks [80] would not be suitable for installation in this space since the required movements do not fit the spatial setting. This example illustrates that the characteristics and requirements of different interaction techniques shape the spatial configuration of an interactive media facade installation and affect whether an interaction technique is applicable in a particular setting at all.

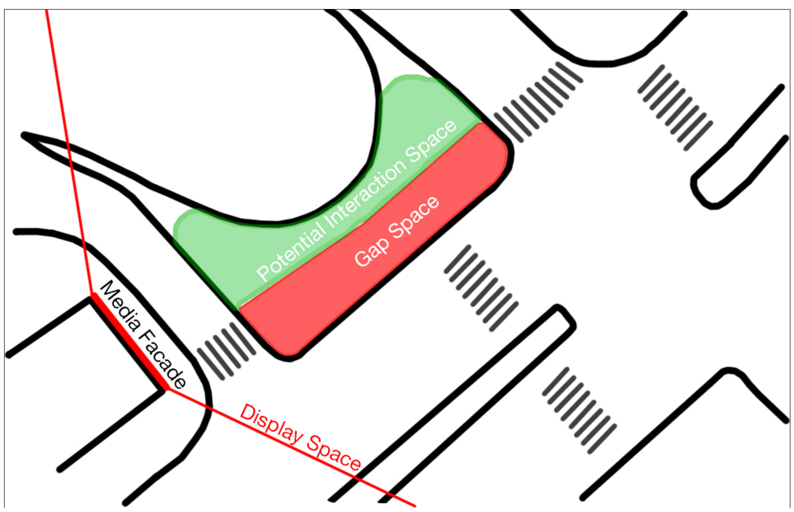


Figure 3.10: The spatial configuration for a camera interface according to Fischer and Hornecker [54]. For a better readability, the streets are not marked as gap spaces.

Having very short interaction distance similarly influences the size and arrangement of the particular spaces. For example, for situated public displays the most common interaction technique is direct touch. The displays are small in comparison and a user can usually get close enough to the display to touch it. For interaction by direct touch, the required minimal interaction distance for a user is minimal, since the user stands most an arm's length away from the display, directly touching it. Hence, the gap space between user and display is minimal as well. In the vast majority of media facade installations, due to their enormous size and required minimal viewing distance, it is not possible to create an interactive media facade installation which users can interact by touching it. In these cases, the media facade installations require a minimal viewing distance, as well as techniques for interacting at a distance to provide means for interactivity. However, as can be seen in the case of the installation *La Vitrine Culturelle* [82], interactive media facade installations can also be created at ground level alongside a building, allowing users to get close to the display when interacting with it (see Section 2.2.2). On the other hand, this can drastically reduce or even remove the gap space between user and media facade, but on the other hand, this also minimizes the potential interaction space and thus also limits the number of potential users.

When designing an interactive media facade installation, designers ideally know from the beginning of the design process which interaction techniques are applicable for the intended purpose in the given space. However, in contrast, adapting a particular interaction technique to make it work in a particular setting is often also an explicit design intention. However, from the aforementioned implications, we can make the following observations as general guidelines to inform the design of interactive media facade installations in terms of choosing appropriate interaction techniques:

Observation 4. *Mobile interfaces are well suited for small and also non-coherent display spaces. They can maximize the size of the potential interaction space.*

Observation 5. *Mobile interfaces engage dynamic social spaces throughout the potential interaction space, where people congregate, being attracted by the system, and have a Shared Encounter.*

Observation 6. *Stationary interfaces are well suited for settings where the display space that could serve as the potential interaction space in front of a media facade is relatively small.*

Observation 7. *Stationary interfaces come at the expense of limiting the number of potential users to a minimum.*

Observation 8. *Stationary interfaces create stationary social spaces around the interaction device.*

Observation 9. *Camera-based interfaces that require whole body interaction might require large, coherent spaces as a potential interaction space. The size of this space depends on the amount of physical body movement involved in the interaction.*

3.4 Summary

In this chapter we investigated the properties and characteristics of interactive media facade installations holistically. We derived taxonomies to categorize interactive media facade installations based on different characteristics and put the installations into context in order to allow for a comparison of different techniques and settings. The extent to which a media facade allows passersby to participate in the installation, the resolution of the utilized input technology and of the media facade, and the properties framing the overall setting could be identified as salient characteristics that can be considered for describing interactive media facade installation. We contribute to the domain of media facades by providing guidelines and an overview of which approaches and technologies work given a particular setting. They can be utilized to both analyze existing installations and to plan and design new ones.

4

Interacting with Media Facades

In this chapter we introduce novel techniques to interact with media facades at a distance. With their design, we address the three significant characteristics identified in Chapter 2: (1) Media facades vary in their technical specifications, such as their resolution, which can easily range from hundreds to millions of pixels. To address this, in Section 4.1 we develop an interaction technique for applying interaction through live video patterns to interact with media facades of arbitrary resolutions. (2) Additionally, existing media facade installations that provide means for interaction often require extensive instrumentation of the potentially interactive area. With the GPS lens, we provide a technique to utilize a common mobile device to control a pointer on the media facade by pointing at it with the mobile device, without requiring an additional instrumentation of the potential interaction space around the facade (see Section 4.2). (3) Finally, as they cover whole building facades, media facades can have non-planar, 3D form factors, different from the common 2D rectangular display shape. Facade Map, an interaction technique presented in Section 4.3, allows for interaction with media facades of arbitrary form factors by providing 2D map representations of the media facade's 3D surface. With this, users can interact with the whole facade — including occluded parts — independently of their current point of view. The interaction techniques presented utilize standard mobile devices and hence do not require the deployment of dedicated input devices. Furthermore, they in particular allow for simultaneous interaction by multiple users with a media facade. After introducing the aforementioned interaction techniques, we classify them according to the taxonomies provided in Section 3.2. The contributions of this chapter have been previously published in [24, 67, 64, 195, 68, 66, 111].

4.1 Interacting with Low-Resolution Facades

4.1.1 Introduction

The size, visibility and large audience of media facades offer a great potential for collaborative interaction of multiple users. Due to the physical properties of media facades, the number of potential users interacting simultaneously can be very large compared to smaller situated public displays. Current approaches for interacting with media facades usually involve controlling pointers on the canvas of the facade [25], as well as pushing content onto the facade by sending messages [146]. However, direct and indirect pointing techniques come with the drawback of restricting the number of simultaneous users as follows:

1. Every user needs their own, distinguishable pointer on the media facade's canvas. Each pointer occludes a small piece of the canvas. The number of pointers increases with the number of users, potentially leading to clutter and large content regions on the facade being covered up by pointers.

2. The more users, the more pointers are displayed on the facade. Hence, finding one's own pointer on the facade becomes increasingly difficult for users. Furthermore, when showing pointers on the facade, the facade needs to have a reasonably high resolution to provide enough pixels per pointer to show distinguishable pointers.

For media facades with low resolution, the number of simultaneously displayable pointers — if pointers can be displayed at all — further decreases. One approach to address these issues is to use an absolute and direct interaction technique such as *interaction through live video* [23], which allows for interaction by direct manipulation through live video with displays of arbitrary resolutions. Multiple users can interact simultaneously without the need to display additional pointers on the facade. However, since this a visual approach that was designed and developed for indoor environments with controlled conditions and regular digital displays with sufficient background lighting, it needs further adaptation when applied outdoors with dynamic weather and lighting conditions.

Within the scope of an interdisciplinary research project, in conjunction with the ARS Electronica[1], the Ludwig-Maximilians-Universität[2] (LMU) and the German Research

[1] http://www.aec.at
[2] http://www.uni-muenchen.de

Figure 4.1: The low-resolution media facade of the ARS Electronica Center in Linz, Austria (Photo: Lois Lammerhuber/Photoagentur Lammerhuber). The surface of the building is equipped with about 40,000 illuminating diodes (red, green, blue), turning the windows into pixels. The framed area (red) providing approximately 230 pixels (10x23) was used for the interaction.

Center for Artificial Intelligence[3] (DFKI), we participated in the 2010 ARS Electronica Festival for digital arts in Linz, Austria. During the festival we presented our application *iRiS*[4] — Immediate Remote Interaction System — to a broad audience on two consecutive days. With iRiS, people could both paint freely on the facade by direct touch input on the mobile device's touch screen, displaying the facade in live video, and solve a puzzle displayed on the facade. The goal of the project was to overcome the limitations of current interaction techniques and allow multiple users to interact simultaneously with the low-resolution media facade of the ARS Electronica Center[5] (see Figure 4.1). As mentioned before, using relative and/or indirect interaction techniques may limit the number of users to the number of distinguishable (i.e. identifiable) pointers on the facade. Due to the low resolution of the media facade — the utilized part of the facade provided approximately 230 pixels (10x23) of about one by three meters each — the number of displayable pointers is further decreased. One approach to overcome this limitation would be to use *world in miniature* representations [176, 199]. Techniques that use such a world in miniature representation overcome this limitation entailing macro attention shifts between both the mobile and the media facade.

[3]http://www.dfki.de
[4]http://www.project-iris.org
[5]The overall facade hosts about 40,000 LEDs embedded into 1087 addressable windows.

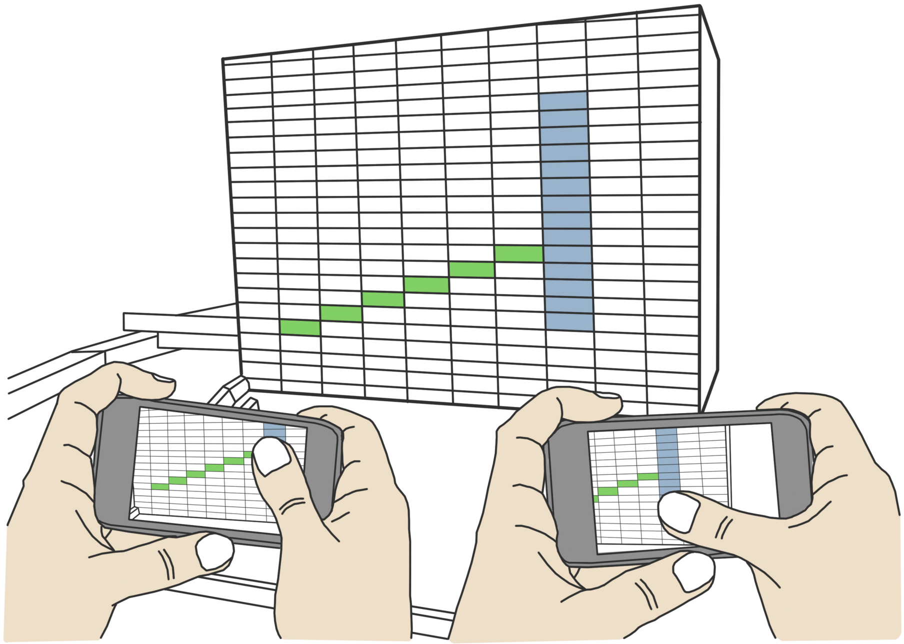


Figure 4.2: Allowing multiple users to simultaneously interact with a media facade through live video on mobile devices. Manipulations by all users on the mobile device are immediately displayed on the facade. The colors denote actions of different users.

To avoid occlusion by displaying pointers, as well as the potential costs of macro attention shifts, we applied and extended the concept of *Touch Projector* [23] (see Section 2.2.3) for use in combination with media facades. The media facade can be observed in live video on the mobile device when aiming the camera at the facade. Users can touch the screen of the mobile device to point *through the display* to virtually touch the surface of the media facade. Since the boundaries of the media facade's canvas are tracked within the live video by using computer vision algorithms (i.e., Canny Edge Detection), touch input occurring on the mobile device's screen can be *projected* onto the facade. With this, we give the user the impression of directly touching the building (see Figure 4.2). To calculate the mapping between touch input on the mobile device and the facade's canvas, the mobile device constantly sends video frames over a wireless data connection to a server which calculates the spatial relationship to the building. The server further handles all touch events received from the mobile device. Because it was initially developed for indoor scenarios with rather static multi-display environments, using the concept of *Touch Projector* in the wild in a dynamic outdoor setting entails several challenges.

4.1.2 Challenges

Technical Challenges

As confirmed by Dalsgaard & Halskov [45], when implementing applications based on visual means for use in outdoor settings, potential shifts in lighting and weather conditions must be considered. As is common for the majority of media facades, the facade of the ARS Electronica Center is only visible below a certain level of daylight. The original version of *Touch Projector*[23] by Boring et al. was designed for use with regular computer screens with sufficient background lighting in controlled and closed environments. In outdoor settings, weather conditions can have a severe impact on the applicability of technologies. Since media facades are based on equipping a building's surface with light emitting elements, the light emitted by the media facade causes reflections on wet ground, especially in the dark. Hence, we needed to substantially change the tracking algorithm to allow for outdoor use. To identify the facade within the live video, we display a white frame around the target facade by lighting the outer pixels permanently (see Figure 4.3). This white frame can be detected with the help of *Touch Projector's* image processing features, such as contrast correction, edge detection and corner detection to find coherent edges. Once the white frame is detected, we use the perspective distortion of the building's outline (the white frame) to calculate the spatial relationship between mobile device and facade. We addressed the issue of reflections on wet surfaces being falsely detected as the facade by making two assumptions:

1. We can assume that users point the mobile devices at the building rather than at the reflection. This causes the reflection to be shown only partly and the frame of the facade to surround the center of focus.

2. The reflections of the facade on the wet ground are slightly jittered. This results in less prominent lines of the building's outline in the reflections.

Early tests on the facade with real users and reflections caused by wet ground around the building showed that these assumptions are sufficient to reliably detect the facade.

In addition to the dynamic environment in which media facades are situated, the media facades themselves impose challenges with their unique features, such as different sizes, form factors, resolutions or optimal viewing distances. As the ARS Electronica Center is situated at the bank of the river Danube, the media facade can be easily viewed

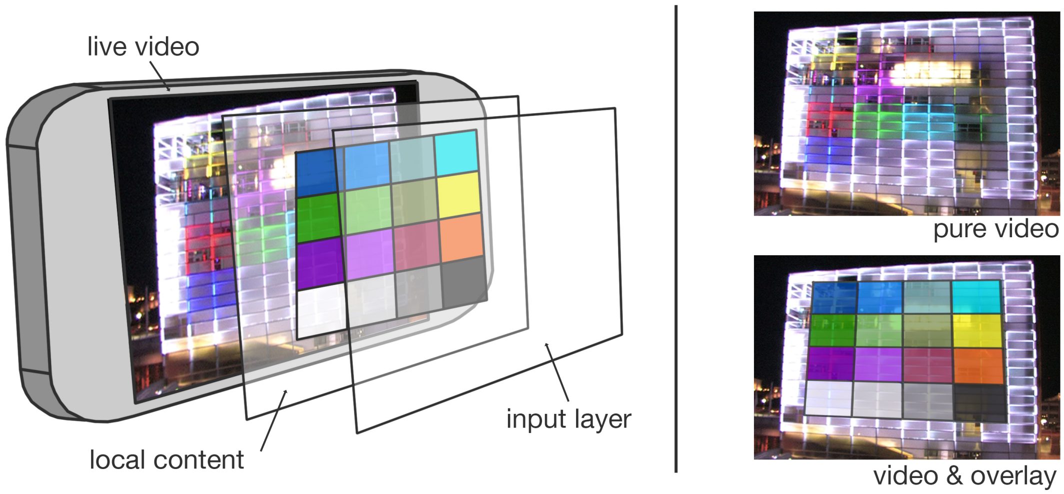


Figure 4.3: Inserting a local content layer in the live video on the mobile device allows for showing the original or augmented live video for each user individually, leaving the canvas of the media facade unaffected.

from the other side of the river, at a distance of more than 300 meters. While enlarging the potential space from which people could interact with the facade and hence the number of potential users, such great viewing distances influence the apparent size of the facade in the live video on a user's mobile device. To address this issue, we made use of the zoom functionality of *Touch Projector*. In the ideal case, the media facade would fit exactly into the live video feed on the mobile device to provide the maximal control display ratio. To achieve this, when the facade is detected within the live video, the zoom level is adjusted automatically on the mobile device to maximize the facade's appearance within the live video image. This ensures a practically constant control display ratio and hence a practically constant resolution factor *Res* for the users, which is independent of their distance from the facade.

Allowing Multi-User Interaction

The enormous size and the great visibility of the media facade allow for a large number of users interacting with it simultaneously. The original version of *Touch Projector* was designed to transform input occurring on the mobile device's screen to the canvas of a remote display. With this approach, temporary feedback and interactive controls are displayed on the display, thus occluding precious screen real estate. In the case of the low-resolution facade of the ARS Electronica Center, this is not an applicable approach.

Interactive controls on the facade waste screen real estate and displaying them decreases the size of the interactive canvas. In addition, the resolution of displayed controls is further limited, due to the low resolution of the facade and the great viewing distances. Furthermore, temporary feedback displayed on the facade (e.g., highlighted regions, tool tips) might interfere with the interaction of other users, since they are usually only relevant for one particular user and the media facade allows for multiple users interacting simultaneously. However, since the live video on the mobile device displays the facade at all times, we can augment the local live video data on the mobile device with controls relevant for the particular local user. Displaying individual feedback directly on the mobile device further avoids the need for macro attention shifts. As depicted in Figure 4.3, the feedback can be displayed by superimposing a personal layer for local content on the live video. This leaves the shared view of the facade's canvas unaffected and displays relevant controls only to the particular user for which they are relevant.

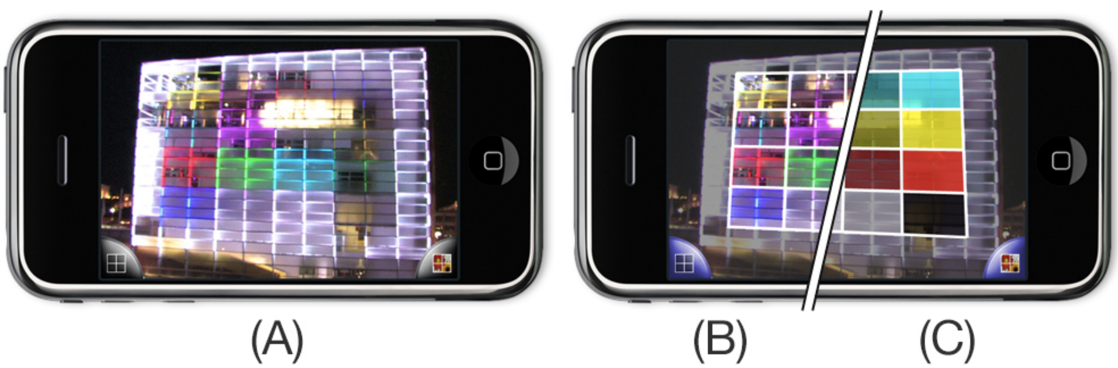


Figure 4.4: An interactive puzzle application displayed on the facade. On request, the original video image (A) is augmented with a grid (B) to determine the tiles or a preview of the solved puzzle (C).

In addition to determining the spatial relationship between the particular mobile device and the facade, the server further stores individual content for each user. The content is stored as image data and can be transferred on request to the user's mobile device. When superimposing individual content on the user's mobile device, the displayed content is automatically distorted by the system for correct alignment with the facade in the live video. The distortion is calculated by applying the inverted transformation matrix (i.e. homography) that was calculated during the process of detecting and tracking the facade's canvas. After the user's individual content is sent to his mobile device as image data, it is displayed as an overlay on the live video image (see Figure 4.4). To handle users' interaction events (i.e., touch input on the mobile device's screen), all interaction events are sent to the server, independent of whether a local item was touched by the user or not. The server can determine and execute the associated actions as it knows the exact locations of all elements. As a result when developing new applications, only interface elements for the mobile client applications have to be designed and the actions they can trigger on the server-side need to be specified. This allows for great flexibility in terms of the heterogeneity found in mobile device platforms. However, since the computation on the server linearly increases with the number of connected mobile devices, this approach also limits scalability.

4.1.3 Implementation & Deployment

To demonstrate the possibility of interacting with media facades through live video on a mobile device as well as the distribution of public and personal content, we provided two applications as a contribution to the ARS Electronica Festival[6] for digital arts. Our deployed prototypes use a dedicated server to control the ARS Electronica Center's media facade through the DMX[7] protocol and to handle the communication with mobile client devices. The server is directly connected to the media facade itself. Similar to Touch Projector, the mobile clients continuously send video frames over a Wifi connection to the server which calculates their spatial relationship to the building by detecting the white frame displayed around the facade. The server further handles all touch events received from the mobile devices.

Solving a puzzle

We demonstrated the distribution of public and personal content with a puzzle application which allowed the users to solve a *15-puzzle* (Gem Puzzle) on the facade. As depicted in Figure 4.4, the game board consists of 15 shuffled tiles and one empty field. The goal of the puzzle is to place the tiles in the original order by making sliding moves which use the empty field. Each tile is represented by eight pixels (i.e., 2 by 4 windows). The users can move a tile by tapping on a tile next to the empty field. The low resolution of the facade did not allow for displaying lines to identify the tiles of the puzzle, which made it difficult for the users to identify them. To allow for identifying the tiles, we

[6] http://www.aec.at/repair/en/
[7] http://en.wikipedia.org/wiki/DMX512

allowed users to superimpose the separating lines on the mobile device (see Figure 4.4). Since each tile is represented by 8 pixels in total it is difficult to determine the correct location of a tile in the overall puzzle. In addition to superimposing lines between the tiles, we allowed people to take a peek at the solved puzzle by requesting a preview image displayed as an overlay on the current appearance of the facade in the live video image (see Figure 4.4). By showing the additional information as overlays in the live video on the mobile device each user can see information only relevant to him and therefore other users are not distracted and we do not waste the screen real estate of the facade.

Figure 4.5: The interface of the painting application. To keep the drawing canvas as large as possible (A), users can switch to the tool palette (B) by performing a sliding gesture.

Painting on the facade

With the second application, we allowed multiple users to paint freely on the facade. Similar to common drawing applications, users can select a tool from a tool palette displayed locally on the mobile device and choose a color with a color picker control. To keep the drawing area as large as possible, tool palette and local controls are initially hidden. They can be activated by performing a sliding gesture next to the live video image on the mobile device's screen. As can be seen in Figure 4.5, the mobile device then locally displays the tool palette. After closing the tool palette (i.e., by performing a sliding gesture in the opposite direction), the user can apply the selected color and tool to the building by touching and dragging on the representation of the facade in the live video. The controls are again displayed locally on the mobile device, since the facade does not have a resolution high enough to display controls at all. A further reason is that the controls are only relevant to one particular user and would interfere with the interaction of other users if displayed on the facade.

User Feedback

During the ARS Electronica Festical, we had the opportunity to present both applications to a broad audience. We handed iPhones with the application already running to users without any further instructions. They observed other users interacting and immediately started interacting with the facade themselves. In the given setting, up to three users were able to interact with the facade at the same time. We ensured that at least two users interacted simultaneously at all times by recruiting users to pitch in when only one user was interacting. The usual way to deploy a mobile device application for interacting with a media facade would be to download the app from an app store or another online source. However, downloading the application was not possible since we had to use a restricted network to access the facade. Furthermore, since we accessed the phone's camera, the application was not allowed in the iOS App Store at that time. While the application itself did not limit the number of users, these restrictions limited the number of simultaneous users to the number of devices we could pass around and oversee. Nevertheless, with three users interacting simultaneously we were able to observe interesting scenarios including collaboration between the users.

We deployed our applications on two consecutive days for 2 hours each day. During both time slots, approximately 50 different users used our applications to interact with the facade. We collected feedback from 15 users (5 female; average age was 26.1) after interacting with facade. The participants were given a 3-minute introduction the applications and following the approach of Burmester at al. [33] for mental note taking, we asked them to keep a *mental note* for every occurrence of positive or negative emotion. After the introduction, each user was given 10 minutes to interact with the facade using the paint application. During that time, they were allowed to pick colors and paint freely on the facade. We video-recorded the participants interacting with the building and we interviewed each of them for 10 minutes immediately after interacting with the facade. We audio-recorded the interviews for analysis. We used an investigative two-step interview process based on the laddering technique [152], deducing qualitative aspects of the interaction based on positive emotions. We started each interview by referring to the *mental notes* which the participants took on positive and negative emotions while interacting with the facade. We asked them why they thought a positive or negative aspect occurred. One user stated that he enjoyed the freedom of *picking, mixing* and *applying* any possible color to the building. He further elaborated: "*Because I can do it completely by myself and it does not happen automatically.*" We received highly positive feedback about the general idea of our prototype. As most of our interviewees had a background in art and/or architecture, the fact that they could change (and observe) the facade in real-time was mentioned positively by nearly all of them. One noteworthy statement about painting on a facade in general was made by one participant, stating: "*Finally, I have control over a building.*"

In terms of interacting in parallel, we found that users liked both collaborative and competitive interaction. When collaborating, the users teamed up to create a joint outcome. When competing, the users usually competed for exclusives where they tried to exclusively display their content on the facade. However, the interaction style strongly depended on them knowing each other beforehand. In terms of collaborative interaction, we observed an interesting interaction style: Two participants used the paint bucket tool to color the whole facade with different colors alternately to create a stroboscope-like effect. On the other hand, simultaneous interaction was also used for disrupting the drawings of others instead. This disturbance was generally observed when these users did not know each other. Especially when two or three users painted simultaneously, they felt that the unpredictability of the outcome led to an interesting piece of art. One stated: "*This is another layer of fun. You can spoil others' drawings and you can draw together. So it's a new way to combine stuff.*" Another user stated: "*It is really nice when you interact with someone else, you can destroy his drawing which is funny.*" A contrasting statement regarding a collaborative use was made by a participant with a background in media design: "*Well, it was good and bad, because it is good in a way to interact in a parallel way if you know the person, you are working together. But if you don't know the person, you are kind of fighting over the pixels and over the space to draw and it's kind of annoying.*" This type of interaction mostly occurred when users were unaware of each other, which turned out to be a general problem of interacting at a distance.

4.1.4 Contributing to the Open Source Community

The adapted *Touch Projector* approach applied to *iRiS* represents a direct interaction approach that can be used with arbitrary rectangular shaped displays. We adapted Touch Projector to outdoor scenarios, creating a universal interaction tool. We detect a white rectangular frame which is displayed around the potentially interactive display. However, there is still the need for a server application to which the mobile device sends the camera frames grabbed for detecting and tracking the display. Hence, the computation is handled by the server and the scalability of the approach for multiple users strongly depends on the available bandwidth. Our approach provides an intuitive way for direct interaction with a distant display (e.g., media facade) in general. For this reason, we contributed a modified — improving the scalability — and generalized version of our interaction client to the TUIO[8] framework. Introduced by Kaltenbrunner et al. [100], the TUIO protocol is the de-facto standard for multitouch applications, utilized to interact with a wide range of interactive digital screens. By using TUIO, our approach can be easily used to enable distant interaction with any multitouch application running on a rectangular display. The only requirement is a Wifi connection between the display and the mobile client device. TUIO defines a generalized protocol for transmitting an abstract description of interactive surfaces, including touch events and tangible object states.

In order to avoid the need for a dedicated server and to decrease the amount of data sent from the mobile device, we need to perform all computations directly on the mobile device and only need to send TUIO messages to the particular display (e.g. facade). The recognition and tracking of the display is the central part of our client. Both are achieved by analyzing the frames of the live video from the mobile device. To make it independent of a server application, we implemented the tracking component of our framework completely on the mobile device. The tracking component detects the rectangular borders of a display that occurs in the live video. There is no need to display a white frame around the display as in the previous version. The applied algorithm utilizes the Open Source Computer Vision Library OpenCV[9] for the analysis of image data. The process of detecting a public display in the live video on the mobile device consists of the following four steps (see Algorithm 1 for pseudocode):

[8]http://www.tuio.org
[9]http://opencv.willowgarage.com

1. The current frame of the live video is converted into a gray-scale image. This basic step is necessary to achieve optimal results. In order to enhance dark structures in the gray-scale image, we apply Erosion[164] as a morphological filter. This sharpens the edges of the display in the image (see Figure 4.6, 1).

2. We apply a Canny Edge Detector[36] to extract the edges of the filtered image. To improve the result and to close gaps between the detected edges, we apply Dilation[164] as an additional morphological filter. This expands bright structures in the filtered image (see Figure 4.6, 2).

3. We apply contour detection to the filtered image. Since the dominant shape of public displays is rectangular, all detected contours are verified accordingly. Contours which form a rectangle are classified as a possible display (see Figure 4.6, 3).

4. To complete the display recognition, the rectangle whose center is nearest to the camera picture center is declared as a display (see Figure 4.6, 4).

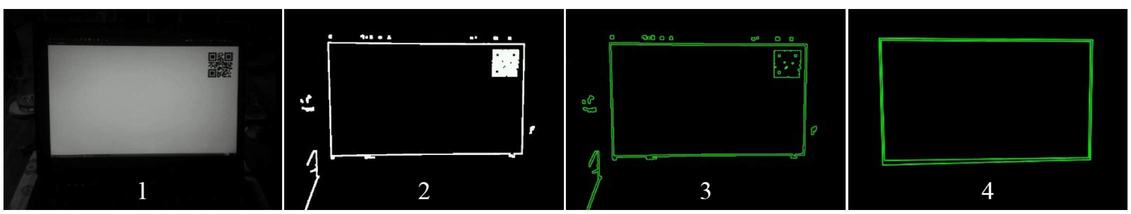


Figure 4.6: Detecting a distant display: (1) Converting the image into a gray-scale image and applying Erosion. (2) Applying Canny Edge Detection and Dilation. (3) Verifying contours. (4) Finding the rectangle nearest to the center of the image.

The algorithm is initially applied to every frame of the live video until the first display is successfully detected. At this point, the display detection is paused to save resources and to speed up the tracking. When moving the mobile device, the display tracking continues to update the position of the detected display in the live video. The motion of the mobile device is detected with the help of the accelerometer. As mentioned before, we utilize the TUIO protocol for sending the touch information from the mobile device to the public display. As can be seen in Figure 4.7, touches are only detected if they are recognized within the borders of the detected display in the live video. Before transmitting the touch information to the display, the touch coordinates need to be mapped on the dimensions of the detected display. Hence, the upper left corner of the rectangle presents the zero point in a two-dimensional coordinate system. As depicted in Figure 4.7, the mapping of the touch coordinates from the screen of the mobile device to coordinates on the display can be achieved simply by subtracting the difference between the zero point of the above explained coordinate system and the zero point of the overall touch area from the touch input coordinates. When touching the display of the mobile device, the display tracking is paused if active. With this approach, a stable computation of the mapping can be achieved. In this case the last detected display frame is used to compute the mapping.

Algorithm 1: Detecting a rectangular display in the live video of the mobile device.

Data: live video stream
Result: screen of the remote display

foreach *video frame* **do**
    convert frame to gray scale();
    Erosion to sharpen edges();
    Canny Edge Detection to extract edges();
    Dilation to expand bright structures();
    contour detection();
    verify rectangles();
    return *rectangle closed to the center of the frame*;
end

By using the TUIO protocol to communicate user input to a display that is not capable of handling touch input, we allow for simple porting of existing multitouch, TUIO capable applications to public displays with a Wifi connection between display and mobile input device as the only requirement. Furthermore, we provide a simple way of enhancing non-interactive public displays with the capability to handle user input.

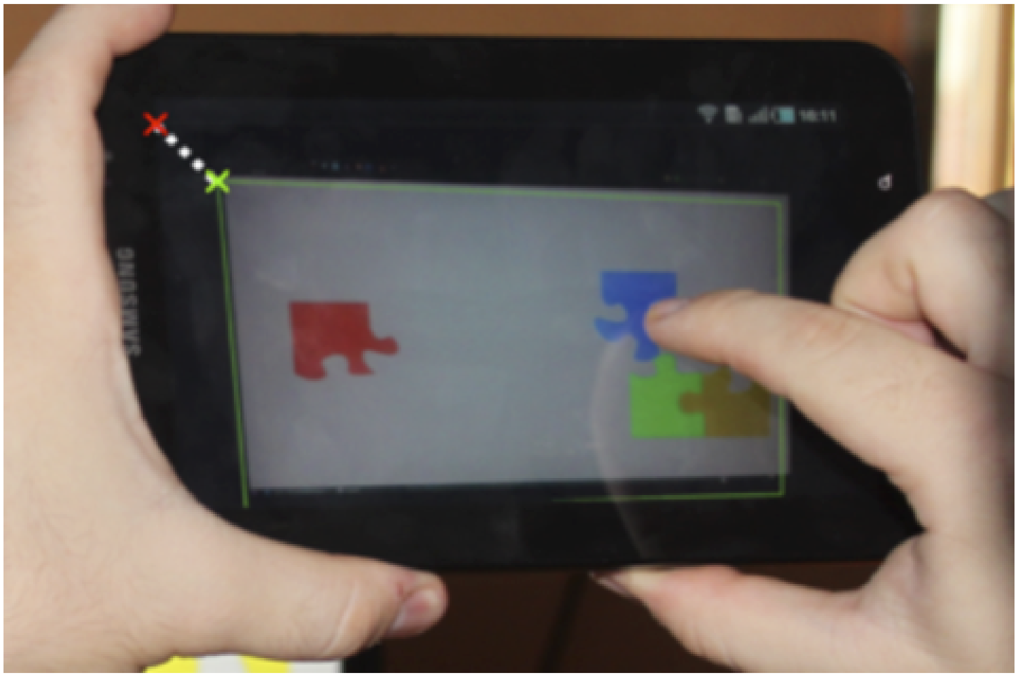


Figure 4.7: A detected display (green rectangle). The red x denotes the zero point of the touch-capable area. The green x denotes the zero point of the coordinate system, which is used for mapping the touches.

4.1.5 Summary & Classification

In this section, we extended the concept of Touch Projector [23] to allow multiple users to simultaneously interact with a media facade through live video on their mobile device. Although reliant upon visual means, our approach can be used under various weather conditions on any digital display that has or can display a white frame. Furthermore, it is also independent of the resolution of the media facade and can thus be applied for arbitrary resolutions. As depicted in Figure 4.8, we utilize the touch screen of a mobile device as a high-resolution input device to interact with media facades of arbitrary resolutions. We superimpose individual content (i.e., UI elements that are not of interest to all users at once) on the live video to save screen real estate on the facade. While this is necessary when addressing low resolution media facades, it constitutes a very general mechanism when many users may interact on larger digital surfaces with their mobile devices: When feedback only affects (or is intended for) a subset of these users, our approach does not distract or disturb others while they interact with the display.

Concerning the general properties of the setup necessary to apply the presented interaction technique, we provide a method to interact with a media facade by direct touch input on a mobile device's screen, resulting in a mobile setting where users can walk around the potential interaction space while interacting. Since the communication of the mobile device and the media facade is carried out over a common Wifi connection and since the interaction technique is based purely on visual means, we do not need any further instrumentation of the environment around the facade. This results in a classification into the taxonomy of media facade interaction (see Section 3.2.4) as depicted in Figure 4.9.

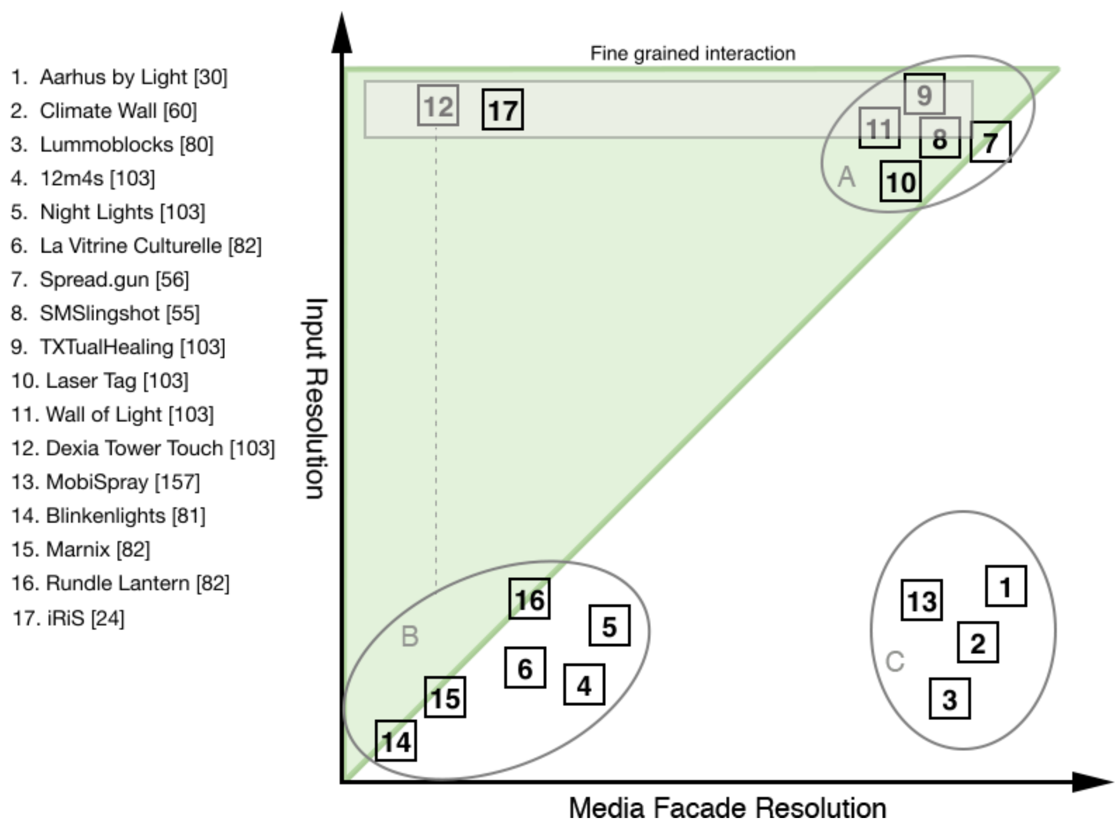


Figure 4.8: Placement of the introduced interaction technique (iRiS) in the classification (consider also Section 3.2.1) by input and media facade resolution (17). Aside from low-resolution media facades, it is also applicable to media facades of various higher resolutions.

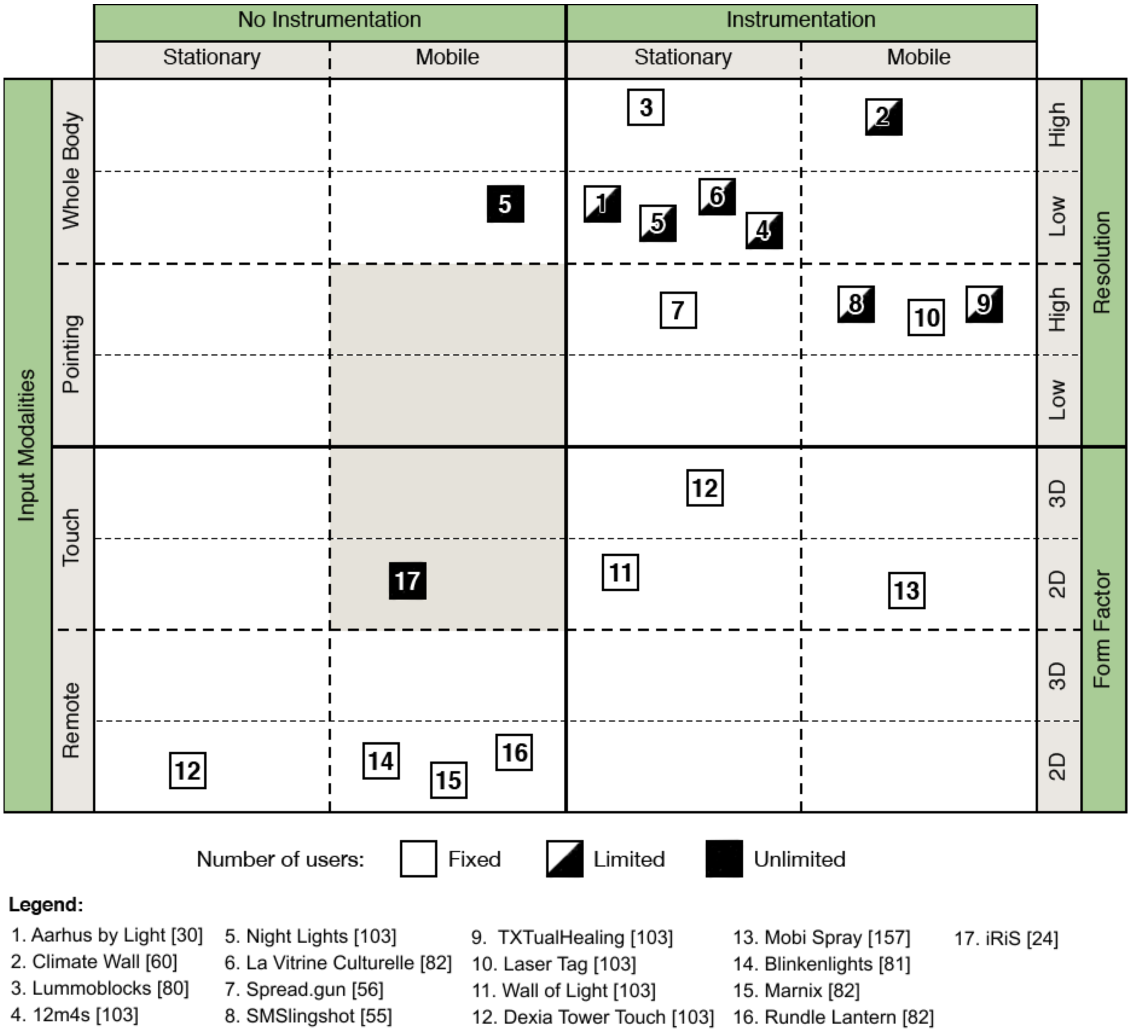


Figure 4.9: Placement of the introduced interaction technique (iRiS) in the classification (see also Section 3.2.4) according the characteristics and the technical specifications of the installation (17).

4.2 Pointing at a Media Facade

4.2.1 Introduction

When developing interactive applications for media facades, one difficulty is the portability of systems and technologies. Interactive installations are usually highly tailored to the specific media facade and its deployment environment. Furthermore, interaction techniques often require the extensive instrumentation of the potentially interactive space with technology (e.g., tracking systems, sensors, etc.) Hence, they usually cannot be transferred to different settings without further effort. However, since media facades usually are situated in outdoor settings, are large in size compared to regular situated public displays, and enclosed by open spaces or plazas, using a positioning system such as GPS[10] can open up ways to provide new portable interaction techniques for media facades. These interaction techniques could be adapted to new settings with ease. With GPS Lens, we introduce a lightweight interaction technique to control pointers on large-scale urban displays like media facades by pointing at the display with common mobile devices (see Figure 4.10). Our approach utilizes the built-in GPS sensor, compass and accelerometer of the mobile device to compute the user's pointing direction. To overcome the limitations of current GPS sensors, we adapt the Magic Lens metaphor introduced by Bier et al. [16, 17] and display a preview of the display's content around the current pointer position on the local screen of the mobile device by applying the *Semantic Snarfing* approach of Myers et al. [132]. With this, we provide a detailed view of the area of interest on the remote display on the mobile device's screen. To overcome the inaccuracy of the current pointer position introduced by inaccuracies of the GPS signal, users can interact with the remote display's content through the detailed view on the mobile device with the high precision of direct touch input.

Pointing in real-world settings

As described in Section 2.2.3, pointing to interact with remote displays is an interaction technique that has been researched extensively. Pointing has been explored in real-world settings to interact with the environment as well as for interacting with distant digital content. In [13], Baus et al. incorporated pointing devices into real-world settings. They introduced ARREAL, an augmented reality outdoor navigation system in which users can use a modified electronic compass as a 3D pointing device to point at buildings in order

[10]http://www.gps.gov

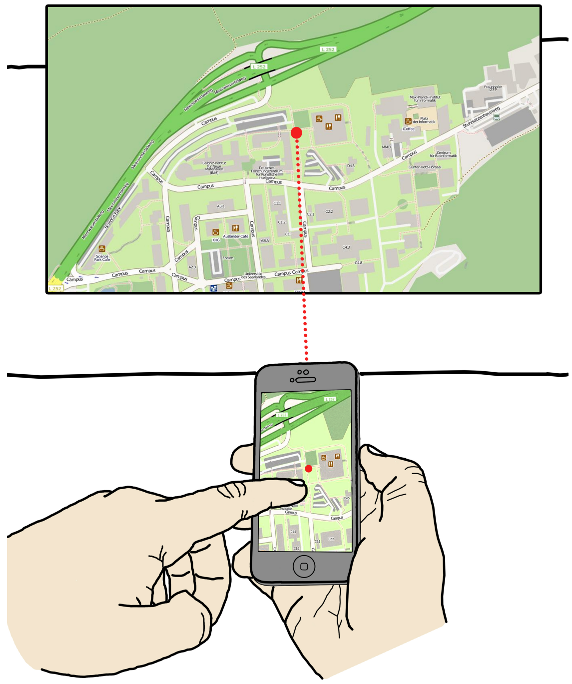


Figure 4.10: Pointing at a display with a mobile device. The content around the current pointer position on the display is displayed as a preview on the mobile device's screen, with which the user can interact by direct touch input. The current pointer position (red dot) is displayed on both the mobile device and the remote display.

to get additional information. However, the described system requires certain instrumentation of the user, as well as detailed spatial information about the environment itself. Simon and Fröhlich introduced a framework to develop mobile geospatial web applications [169], allowing users to point with a mobile device to perform spatial queries based on visibility and field of view. The queries return information on points of interest (POIs) in this direction. Similarly, Lei and Coulton presented the gesture controlled Geo-wand [112]. Users can point with a common mobile device at directions to get information on POIs. In addition, they can manually add information such as photos to the POIs. Both systems provide information on the set of POIs in a certain area. In contrast to this approach, Beer describes GeoPointer [14], a system for detecting a particular building a user is pointing at with a mobile device.

We pick up on these ideas of pointing at objects to get information about them and utilize GPS for controlling virtual pointers on urban screens. In order to be independent from changes in environmental conditions like lighting conditions and to provide a

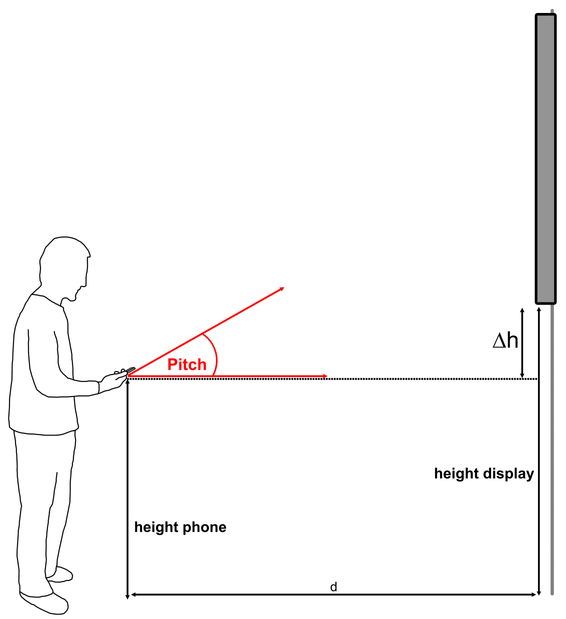


Figure 4.11: Calculating the vertical intersection of the current pointing direction and the vertical display area.

lightweight, low-bandwidth technique to control a pointer on large-scale urban displays, such as media facades, we decided to utilize the built-in GPS sensor, accelerometer and compass of a mobile device in order to compute a user's pointing direction.

4.2.2 GPS Lens

The goal of GPS Lens is to utilize the mobile device of a user as an input device to continuously control a pointer on a media facade. The pointer movement is calculated based on data obtained from the built-in GPS sensor, accelerometer and compass of the mobile device. To point with the mobile device, we assume a posture as described in Figure 4.11, where the user is holding the mobile device in front of his body, at a height that is comfortable for reading the content of the mobile device's display. We consider this to be the optimal posture since the user needs to be able (1) to point with the mobile device at the display and (2) to see the content displayed on the mobile device's screen and to be able to interact with it at the same time. Since GPS only provides ground truth data, we incorporate the height of the phone — the height at which the user holds it — into the calculations.

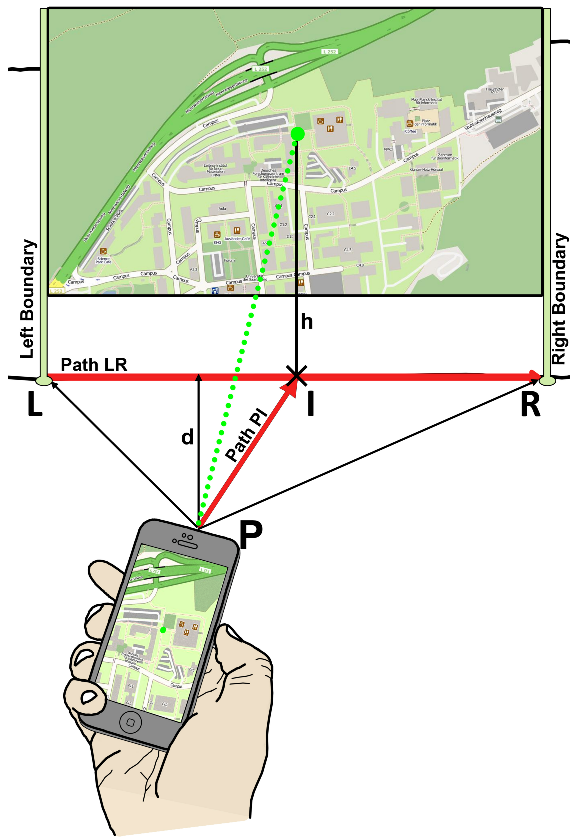


Figure 4.12: Calculating the horizontal intersection of the current pointing direction PI and the display area.

Calculating the cursor position

To control a pointer on a large-scale urban display such as a media facade based on GPS data, we need (1) the current location and orientation of the mobile device, as well as (2) the GPS coordinates of the media facade's boundaries. To map the pointer onto the facade, we need the GPS coordinates of the left and right boundaries, as well as information on the exact height of the media facade, the lower and upper boundaries. The accuracy of the altitude, measured by low-cost GPS receivers, is specified by the manufacturers to be +/- $15m$ in 95% of the measurements. However, utilizing digital elevation models (DEM) [114] to get the altitude of the current user location could easily circumvent this inaccuracy. Using the measured altitude for calculating the pointer position would introduce an error which could drastically reduce the pointing accuracy and might make the system unusable. When interacting with a media facade by pointing at it, we can assume that the user is standing in front of the actual media facade in a direct line of sight. For the sake of simplicity, we assume that the area in front of the facade is flat and we can calculate a relative difference in altitude between mobile device and facade. If there was a significant pitch in the area in front of the facade, the difference in altitude could be easily measured in advance and mapped to the relative position of the user. In terms of horizontal accuracy, although common mobile devices state an accuracy of $5m$ as an upper bound, our measurements showed that the actual accuracy is often better. For optimal conditions, we measured an average accuracy of 1 - $2m$ for an iPhone 5, compared to dedicated GPS sensors. Such an accuracy is achieved on current mobile devices by incorporating additional information such as available Wifi networks, the current cell ID or the last known position. When computing a path between two locations on the curved surface of the Earth, for a sufficiently long distance between two locations, we need to incorporate the curvature of the Earth. In the case of pointing at a media facade, we are computing the path from the pointing device (e.g., the user's mobile device) to the target point at the facade. Since we can assume that a user stands relatively close to the media facade while interacting — at most a few hundred meters away — we can apply simplifications when computing the pointing direction and we do not have to incorporate the curvature of the Earth in our calculations.

All formulas applied in our calculations are standard formulas to calculate geographical relations between latitude-longitude points [79, 170]. As depicted in Figure 4.12, the pointing direction is calculated as follows: The user is pointing at the display if the pointing direction of the mobile device in the user's hand intersects with the media facade's boundaries both horizontally and vertically. In general, when for paths between geographical coordinates, we have to distinguish between *heading* and *bearing*. The bearing — which is also referred to as the forward azimuth — is the angle between North and the destination while the heading — which varies when following a path on the Earth's curved surface — is the angle between North and the moving direction when moving on a path on the Earth's curved surface. In the description of the calculations, since we are calculating at short distances of several meters, we can use the term heading to describe directions and orientations. There is no need to distinguish between heading and bearing because for sufficiently short distances, as in our case, both can be considered equal. To detect a horizontal intersection, we initially calculate the heading from the location (P) of the pointing device to the left (L) and right (R) display boundaries. Afterwards, the heading of the current pointing direction from P is calculated. If it is between the heading from P to L and P to R, the pointing ray cast by the mobile device intersects the area between the horizontal display boundaries. In general, the geographical intersection point I can be calculated as:

$$I = intersection(path(L, R), path(P, currentHeading)) \tag{4.1}$$

Since we have left and right display boundaries in geographical coordinates, the GPS coordinates of the mobile device's locations and its orientation, the geographical coordinates of the intersection point of the pointing direction with the display I can be calculated as follows:

At first, we calculate the *bearing* — the current heading when traveling on a great-circle arc — from the mobile device's location to the left and right display boundaries L and R. The bearing only has to be re-calculated for every change of the mobile device's location. In general, the bearing from a point $p_1(\phi_1,\lambda_1)$ to a point $p_2(\phi_2,\lambda_2)$ — where ϕ denotes the latitude and λ the longitude — can be calculated as

$$\theta = \arctan 2\,(\sin(\Delta\lambda) * \cos(\phi_2), \cos(\phi_1) * \sin(\phi_2) - \sin(\phi_1) * \cos(\phi_2) * \cos(\Delta\lambda)) \tag{4.2}$$

Second, for each change in the mobile device's heading, we check if the current heading of the mobile device is between the bearing from the mobile device's location P to the left and right display boundaries L and R. If this is the case, the pointing ray cast by the mobile device horizontally hits the display. Hence, after detecting an intersection between the pointing ray and the horizontal boundaries of the facade, we then — and only then — need to check for an intersection with its vertical boundaries to make sure the pointer is within the visible area of the media facade (see Figure 4.11). To do this, we first calculate the intersection point $I(\phi_3, \lambda_3)$ of the pointing direction and the horizontal display line. Therefore, we need the point L and the bearing from L to R — or the point R and the bearing from R to L — and the mobile device's location P, as well as its current heading. Given two points $p_1(\phi_1,\lambda_1)$ and $p_2(\phi_2,\lambda_2)$, and the respective bearings θ_1 and θ_2, we can calculate the coordinates of the intersection point as:

$$d_{1,2} = 2 * \arcsin\left(\sqrt{(\sin^2(\Delta\phi/2) + \cos(\phi_1) * \cos(\phi_2) * \sin^2(\Delta\lambda/2))}\right) \tag{4.3}$$

$$\phi_1 = \arccos\left(\sin(\phi_2) - \sin(\phi_1) * \cos(d_{1,2}) / \sin(d_{1,2}) * \cos(\phi_1)\right) \tag{4.4}$$

$$\phi_2 = \arccos\left(\sin(\phi_1) - \sin(\phi_2) * \cos(d_{1,2}) / \sin(d_{1,2}) * \cos(\phi_2)\right) \tag{4.5}$$

if $\sin(\lambda_2 - \lambda_1) \geq 0$

$$\theta_{1,2} = \phi_1 \tag{4.6}$$

$$\theta_{2,1} = 2 * \pi - \phi_2 \tag{4.7}$$

else

$$\theta_{1,2} = 2 * \pi - \phi_1 \tag{4.8}$$

$$\theta_{2,1} = \phi_2 \tag{4.9}$$

$$\alpha_1 = (\theta_1 - \theta_{1,2} + \pi)\,\%2 * \pi - \pi \tag{4.10}$$

$$\alpha_2 = (\theta_{2,1} - \theta_2 + \pi)\,\%2 * \pi - \pi \tag{4.11}$$

$$\alpha_3 = \arccos\left(-\cos(\alpha_1) * \cos(\alpha_2) + \sin(\alpha_1) * \sin(\alpha_2) * \cos(d_{1,2})\right) \tag{4.12}$$

$$d_{1,3} = \arctan 2\left(\sin(d_{1,2}) * \sin(\alpha_1) * sin(\alpha_2), \cos(\alpha_2) + \cos(\alpha_1) * \cos(\alpha_3)\right) \tag{4.13}$$

$$\phi_3 = \arcsin\left(\sin(\phi_1) * \cos(d_{1,3}) + \cos(\phi_1) * \sin(d_{1,3}) * \cos(\theta_1)\right) \tag{4.14}$$

$$\Delta\lambda_{1,3} = \arctan 2\left(\sin(\theta_1) * \sin(d_{1,3}) * \cos(\phi_1), \cos(d_{1,3}) - \sin(\phi_1) * \sin(\phi_3)\right) \tag{4.15}$$

$$\lambda_3 = (\lambda_1 + \Delta\lambda_{1,3} + \pi)\,\%2 * \pi - \pi \tag{4.16}$$

In order to calculate the height of the vertical intersection in meters, we first calculate the distance d between the mobile device's location P and the horizontal intersection point I on the path LR. In the geographical context, this would mean calculating the Haversine [79, 170] distance. Since we are dealing with a sufficiently short distance, we can use the less computation-intensive Pythagorian theorem. Along meridians, the Pythagorian theorem does not introduce errors. Otherwise, the errors depend on distance, bearing and latitude. However, they are negligibly small for short distances. Hence, given the Earth's mean radius $R = 6371km$ we compute the distance $d_{P,LR}$ as

$$x = \Delta\lambda * \cos(\phi_P) \tag{4.17}$$

$$y = \Delta\phi \tag{4.18}$$

$$d_{P,LR} = R * \sqrt{x^2 + y^2} \tag{4.19}$$

With the distance $d_{P,LR}$, the current pitch — measured with the built-in accelerometer — of the mobile device and the height difference between mobile device and lower display boundaries, we can calculate the height of the vertical intersection in meters as follows:

$$h = d_{P,LR} * 1000 * tan(currentPitch) - \Delta height \tag{4.20}$$

To transfer the geographical coordinates and the height of the intersection into display coordinates on the media facade, we need to create a mapping between both coordinate spaces. Therefore, we map the calculated horizontal and vertical values of the geographical coordinates to display coordinates relative to the upper left corner of the media facade — starting from (0,0) — with respect to the media facade's resolution. Hence, the coordinates are generally mapped as follows:

$$(x, y) = diplaySizePixel/displaySizeMeters * hitpoint \tag{4.21}$$

While moving the mobile device, the pointing direction and the target point on the media facade's screen are continuously calculated as listed in Algorithm 2:

Algorithm 2: Calculating the display coordinates when pointing at a media facade with a mobile device.

```
Data: source location (P), left display boundary (L), right display boundary (P),
      orientation of the mobile device, heading of mobile device
Result: display coordinates x, y or null
while pointing do
    calculate bearing(P,L) ;                                  // Equation (4.2)
    calculate bearing(P,R) ;                                  // Equation (4.2)
    // Check if horizontal display boundaries are hit
    if bearing(P,L) < heading of mobile device < bearing(P,R) then
        calculate horizontal intersection point I ;  // Equations (4.14), (4.16)
        calculate distance(P,I) ;                             // Equation (4.19)
        calculate pointing height h ;                         // Equation (4.20)
        // check if vertical display boundaries are hit
        if lower display boundary < h < upper display boundary then
            calculate display coordinates x, y ;              // Equation (4.21)
            return x,y;
        end
    end
    return null;
end
```

4.2.3 Managing Jitter

The goal of our work is to provide a system that utilizes data from the built-in sensors of a mobile device to allow interaction with large-scale urban displays through manual pointing. To realize this approach, we have to address additional challenges:

1. When controlling a pointer on a remote display using manual pointing, the natural tremor of the human hand introduces jitter to the pointer which reduces the accuracy of the interaction. This jitter naturally increases with increased pointing

distance to the display. Due to their large size and subsequent great visibility, this is especially an issue for large-scale urban displays such as media facades.

2. The inaccuracy of low-cost GPS sensors influences — and might decrease — the accuracy of pointing and it might introduce an offset as well. For example, a horizontal shift of the cursor position might appear when moving parallel to the media facade while continuously pointing at a fixed point.

3. Furthermore, with increased interaction distance, the sensitivity of the pointer increases. The same movement angle of the pointing device causes a larger movement of the pointer on the display.

We address these issues through the following means: In order to counterbalance the introduced jitter, we introduced movement thresholds and smooth the pointer movements over a time interval Δd. To smooth the movements, we apply the mean value of the movement [107], computing the unweighted mean of n successive data points as follows:

$$m = \frac{1}{n} * \sum_{i=1}^{n} x_i \tag{4.22}$$

not only do we use the mobile device as a pointing device, we also utilize the mobile device's screen to display a preview buffer of the media facade's content around the calculated cursor location (see Figure 4.13) following the Semantic Snarfing approach by Myers et al [132]. They showed that this technique is significantly faster and less error prone than direct manual pointing. This preview buffer serves as a detail view of the targeted area through which the user can interact with the media facade's content by direct touch input on the mobile device's screen. The touch input is mapped and transferred to the actual facade. The size of the area around the pointer which is displayed in the preview buffer can be increased with the interaction distance or a decrease in the accuracy of the GPS signal. Through this, we can counterbalance a decreasing accuracy of the pointing as well as an increased sensitivity of the pointer movement for increasing distance. GPS Lens offers the user the option to temporarily freeze the content preview on the mobile device by pressing a button. The frozen image allows the user to interact with the displayed content without any jitter at all. Hence, the user can interact with higher precision. By pressing the button again, the user can re-activate the pointing mode.

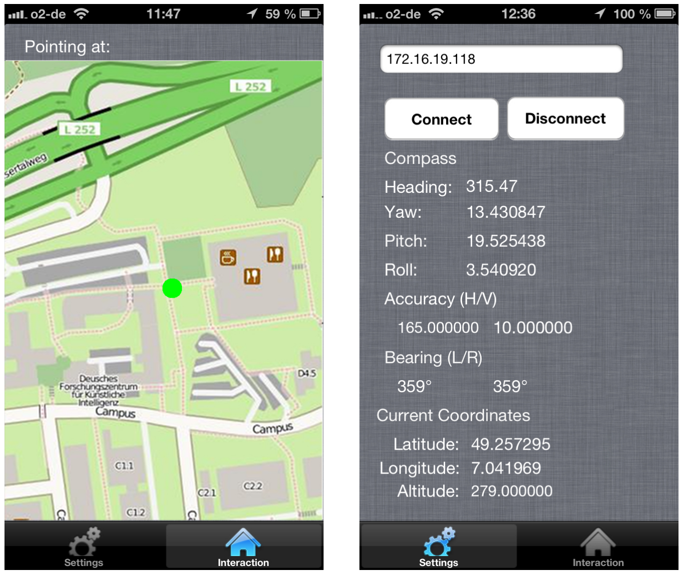


Figure 4.13: The interface of the mobile client prototype. Left: The detail view displaying the content around the current cursor position. Right: A settings view, displaying the current sensor data, utilized for calculating the cursor position.

4.2.4 Implementation

The implementation is built upon a common client-server architecture where the application running on the remote display is listening for standard TUIO messages. We implemented the prototype of our client application for both the Apple iOS and Google's Android platform, utilizing iPhone 5 and Nexus 4 devices as the mobile pointing devices. The mobile client is connected over a constant Wifi connection to an application running on the facade. All calculations to estimate the pointing direction are performed locally on the mobile device itself. The target location of the pointing is continuously calculated on the mobile device for updates of the location or orientation of the mobile device. After mapping the intersection point of the cast pointing ray with the remote display, the mobile device sends the display coordinates combined with a user ID and a zoom level indicator to set the size of the content detail view over the Wifi connection to the server application on the remote screen. The remote screen continuously updates the cursor position and sends back the content for the content detail view back to the mobile device. The remote screen's boundaries as GPS coordinates, as well as its pixel size, are measured in advance and stored as presets on the mobile device.

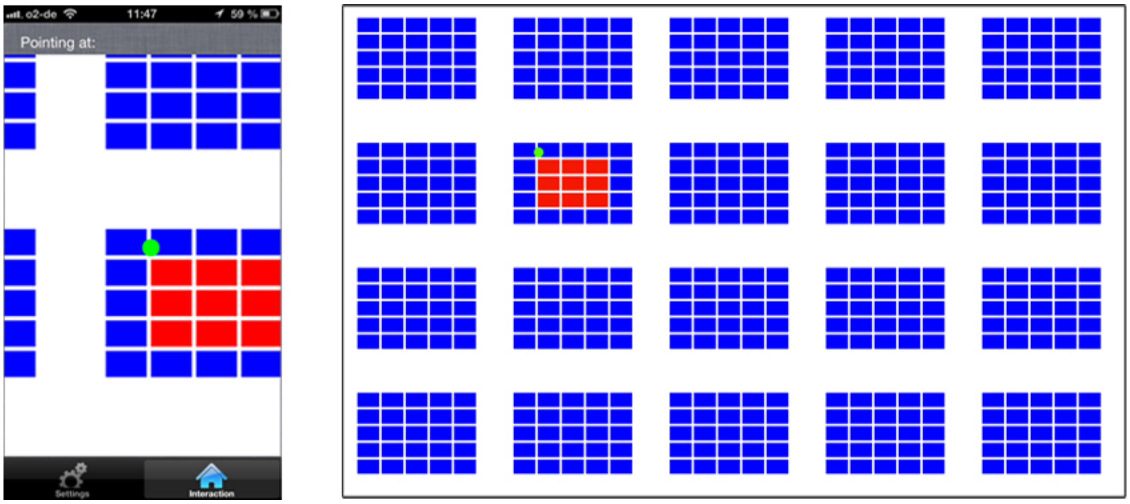


Figure 4.14: Right: The application displaying the targets and the current cursor position on the remote display. Left: The mobile application showing the detail view of the content around the cursor.

4.2.5 User Evaluation and Discussion

To evaluate the accuracy and applicability of the proposed technique we conducted an initial user study with 6 participants (3 male, 3 female) with an average age of 31 years, where we focused on qualitative aspects of the interaction. In particular, we were interested in the question of whether such a technique is applicable to control a pointer on a media facade despite the given inaccuracy of the GPS signal. As depicted in Figure 4.14, each participant had to select 20 randomly chosen blocks out of a grid structure displayed on a projected media facade with a size of 10x8m. The study took place in an outdoor plaza on our university campus under the open sky in order to avoid shadowing of the GPS signal. The participants interacted from a fixed position at a distance of 15m, with an average GPS accuracy of at least 2m which resulted in an average jitter of the controlled pointer of around +/- 29cm. To interact with the projected display, and to select the targets, the participants used GPS Lens running on an iPhone5. To minimize the possible delay that might be introduced by the network connection we used a dedicated wireless LAN network to transmit the data. After completing the tasks we asked the participants to fill out questionnaires about the accuracy of the pointing as well as their experiences while using the system. For each question the participants were asked to rate on a 5-point Likert scale.

The accuracy of the pointing was rated as fairly good (3.17 on a scale from 1 = very inaccurate to 5 = very accurate). Asked about the rating, the participants stated that they were well aware that there were inacurracies while interacting, but this did not decrease the usability of the approach. The participants further stated that due to the content preview buffer shown on the mobile device, the inaccuracy of the pointing had only little influence on their interaction (2.17 on a scale from 1 = did not influence at all to 5 = strongly influenced). Asked about the ease of use, the participants rated the system as easy to use (4.3 on a scale from 1 = very difficult to 5 = very easy). They further claimed not to feel uncomfortable while interacting (1.6 on a scale from 1 = very comfortable to 5 = very uncomfortable). In general, the system was rated positively throughout. Hence, despite the inaccuracy of the GPS signal, GPS Lens can be an applicable technique to control pointers on large-scale media facades in urban outdoor settings. Winkler et al. showed that for a very similar setting the human eye-hand coordination can counterbalance minor errors in the pointing accuracy and allow for a fluent interaction [200].

Discussion

The pointing accuracy of the proposed system strongly relies on the quality of the received GPS signal. Besides their lower accuracy in general, for low cost GPS sensors such as the ones built into mobile devices the signal quality and therefore the location accuracy might be even further reduced by changes in environmental conditions, such as the weather. The quality of the GPS signal therefore remains a critical issue for the pointing accuracy and the applicability of the proposed system. Although we have introduced mechanisms to counterbalance this issue, the accuracy of the location — especially the altitude — needs to be further addressed. However, errors in the accuracy of the pointing also increase with the distance from which a user is interacting. Due to their large size and subsequent great visibility, media facades also offer the potential to interact from great distances. This raises the need to further investigate means to counterbalance inaccuracy introduced by the interaction distance and to further investigate the influence of the interaction distance at all.

For the sake of simplicity, introducing GPS Lens we only considered one user at a time. Nevertheless, the system also supports multiple users interacting in parallel. Users are internally distinguished by a user ID which is enclosed in every message sent from the

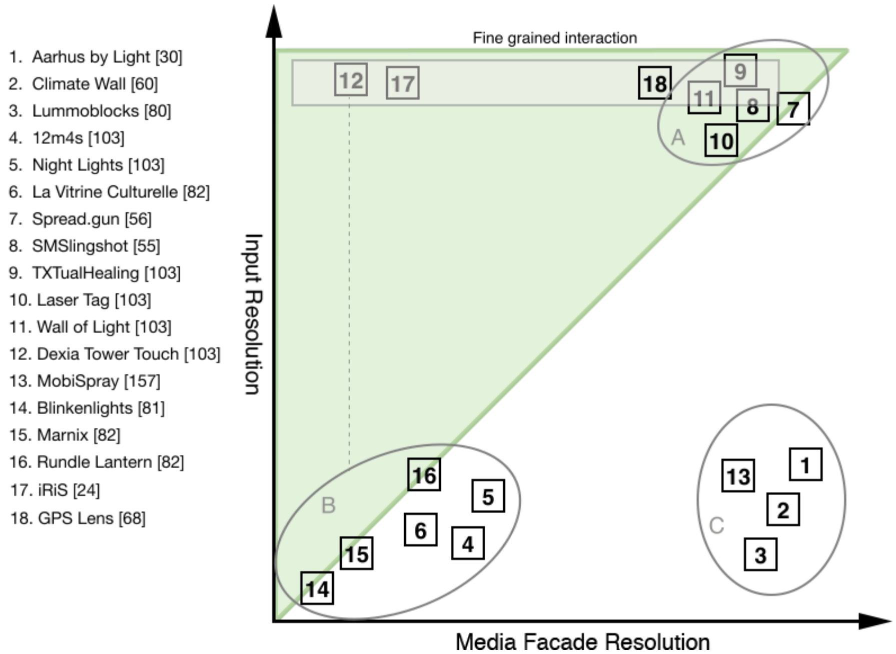


Figure 4.15: Placement of the introduced interaction technique (GPS Lens) in the classification (see also Section 3.2.1) by input and media facade resolution (18). Although initially developed for high-resolution facades, it is also applicable to various resolutions.

mobile device to the application running on the facades. Cursors for different users are displayed in different colors. Since the cursors are displayed on both the facades and the mobile device, each user can easily find the cursor belonging to him. Since each cursor occupies screen real estate, this approach is only applicable on media facades with a sufficiently high resolution. The number of cursors that can be displayed without occluding too much content therefore limits the number of parallel users.

4.2.6 Summary & Classification

In this section, we introduced an interaction technique to control a pointer on large-scale urban displays such as media facades by pointing at them with a mobile device. By using a mobile device as a pointing device, we allow for fine-grained input with high resolution. Since the options for displaying visual feedback on the facade itself while interacting are limited depending on the resolution of the media facade, we utilize the mobile device's screen to display visual feedback locally on the high-resolution screen. Hence, the interaction techniques introduced can be applied to media facades of various resolutions. Figure 4.15 depicts the placement of the GPS Lens in the classification presented in Section 3.2.1. Concerning the general properties of the setup necessary to apply the presented interaction technique, we utilized the built-in GPS sensor, compass and accelerometer of the mobile device to compute the pointing direction of the user. To overcome the limitations of current GPS sensors, we introduced a content preview of the media facade's content around the current pointer position on the local screen of the mobile device. The communication between mobile device and media facade to provide content for the local preview and to communicate interaction data is carried out over a common Wifi connection. Hence, we do not need any further instrumentation of the environment around the facade. This results in a positioning in the taxonomy of media facade interaction (see also Section 3.2.4) as depicted in Figure 4.16.

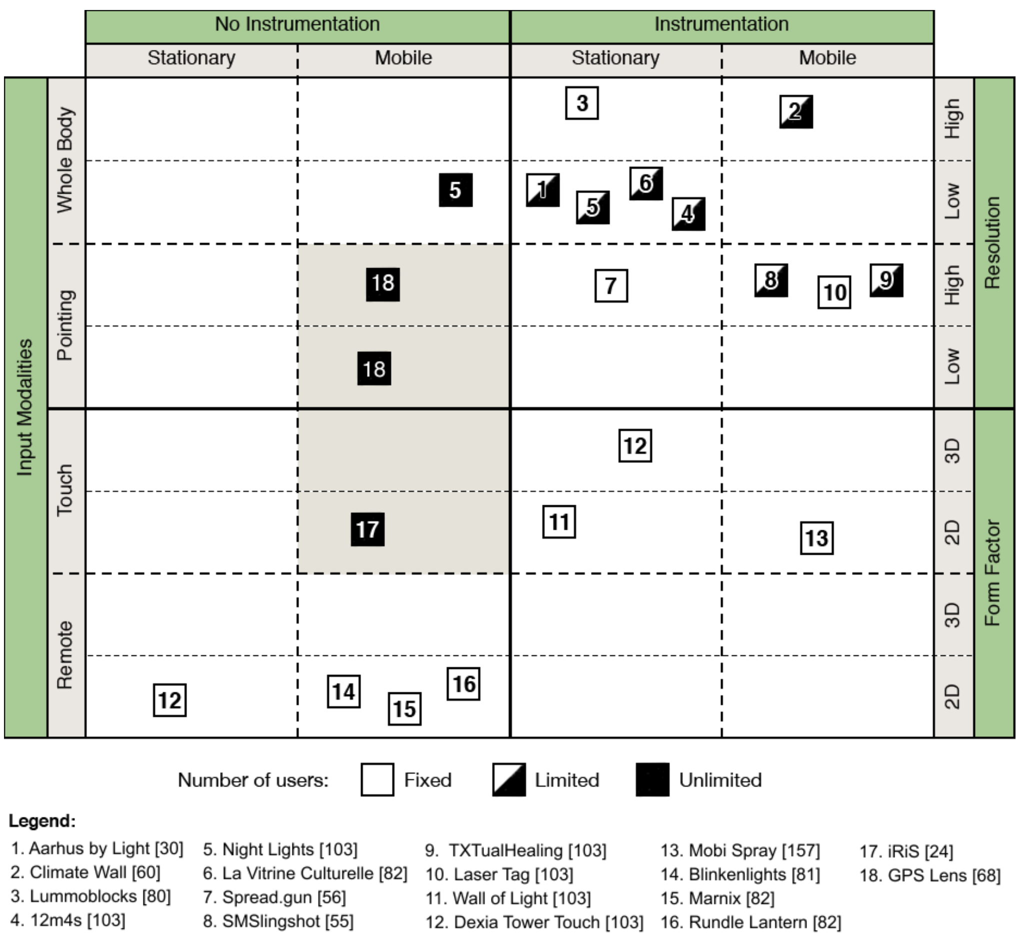


Figure 4.16: Placement of the introduced interaction technique (GPS Lens) in the classification (see also Section 3.2.4) based on characteristics and technical specifications of the installation (18).

4.3 Media Facades with Arbitrary Form Factors

4.3.1 Introduction

As mentioned in the introduction of this thesis, media facades are very large in size. In the majority of cases, they continuously cover more than one side of a building's facade and in some some cases also the roof of a building, which gives the media facade a three-dimensional (3D), non-planar form factor (see Figure 1.3). For the development of interactive installations for media facades, a media facade with a 3D form factor introduces further challenges for the developer. It is very likely that parts of the facade are occluded from certain points of view which hinders continuous interaction with the whole media facade and therefore lowers the quality of interaction.

Describing key challenges for the design of media facade installations, Dalsgaard et al. state that media facades need to be integrated into the physical structures and surroundings of architectural buildings [45]. This often leads to media facades that cover more than one side of a building. In this case, a potential user can only see the part of the facade that is visible from his point of view. Parts of the facade might be occluded although the whole media facade offers a potentially interactive area. Since the majority of interactive installations for media facades require a direct line of sight between the interacting user and the facade, a user can usually only interact with the part of the whole media facade currently visible from the user's point of view (see Figure 4.17). In order to interact with the occluded parts, the user has to re-position himself and move around the building until he can see the remaining parts. Due to the large size of media facades and the spatial settings around the building hosting the facade, moving around can require the user to walk quite a distance, hindering a fluent and continuous interaction and further reducing the quality of the interaction. In the case of MobiSpray [157], the user is required to navigate through a miniature representation of the environment to select a particular part of the facade. Furthermore, the spatial setting might be different around the facade, such that certain techniques for interacting at a distance might not work for all parts of the facade. This could be the case for example when the applied interaction technique requires a particular minimal interaction distance that cannot be maintained. To exploit the full potential and the capabilities of media facades with 3D form factors, the ultimate goal is to allow for a fluent, continuous interaction with the whole media facade, independent of the current point of view. When a user is within the display space around the media facade, the interaction should be independent of the current perspective and point of view. One way to achieve this is to utilize mobile devices as input devices. As a further advantage, users can utilize their own mobile devices and there is no need to provide dedicated or custom-built devices. We need an interaction technique that allows the user to view and access all parts of the facade, including the parts that are not visible from within the current point of view of the user. As one possible way to make the whole media facade accessible for user input, we propose making all parts of the facade visible to the user at all times — as on a cartographic map — on his mobile device and making them accessible by allowing direct touch input on the visualization. We apply cartographic map projections — mathematical algorithms to map coordinates from a 3D into a 2D space — to create 2D map representations of media facades with various form factors. The 2D map representation, which we will call *facade map* from now on, can be displayed on mobile devices such that touch input on the map of the media facade on the mobile device can be directly transferred to the real media facade. Our goal is to provide a set of rules for how to use a 2D map representation of a media facade to allow for continuous interaction with all interactive areas, independent of the current point of view of the user for media facades with various form factors. Furthermore, while interacting with the media facade, we want to create a smooth transition when interacting *over the edge*, where the target area of the interaction continuously moves from within the current field of view of the user to adjacent areas of the media facade outside the user's current field of view (e.g., parts that are occluded from the user's current point of view). This can be seen in Figure 4.17.

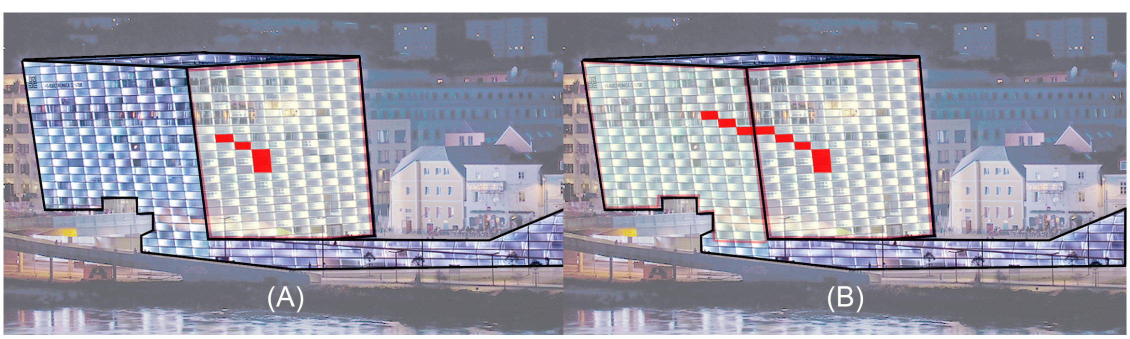


Figure 4.17: (A) Interaction limited to a fixed frame, here one side of the building. (B) Continuous interaction over the edge. (Original photo: Lois Lammerhuber/Photoagentur Lammerhuber)

4.3.2 Cartographic Map Projections

Creating a 2D representation of a 3D surface is one of the key concepts in the area of cartography. In our case, this corresponds to creating a 2D map representation of the media facade's surface. So-called *map projections* are applied to create the 2D cartographic maps that are known and used on a regular basis for navigation and orientation. In [170], Skupin describes cartographic perspectives on information visualization. He discusses how geographic and cartographic approaches can influence the design of visualizations for textual information spaces. The use of map projections and map design are two ideas that we want to adapt to create a 2D representation of a media facade with a 3D form factor, similar to a geographical map, which is basically a 2D representation of the 3D surface of the Earth.

In general, map projections concern the field of mathematical cartography. Map projections denote methods for mapping the dimensions, the shape and the features of the Earth onto a 2D surface, a map. In [183], Tyner introduces the principles of designing maps in general. She gives a general introduction on how to create different maps for a wide range of purposes and she describes which map projections are most suitable to create the particular map, as well as how to apply them. Lev et al. focus more on the theory and spatial aspects of map projections [113]. They describe a wide range of map projection algorithms and how to apply them to create maps for various purposes. Greenhood, on the other hand, describes basic theory about maps [79]. He gives an overview of how to use maps in general, how to orient on a map, and how to read maps to get a variety of information out of them. He introduces the concepts of different coordinate systems, scale, direction and topography of maps. The concept of creating maps as abstract representations of scale in general has been used by humans for more than 6000 years [116]. In [171], Snyder reviews the evolution from the early beginnings of historical maps to currently used map projections. We describe common map projections in detail when we introduce how to apply cartographic map projections to create 2D facade maps.

When mapping the 3D surface of the Earth onto a planar 2D surface such as a cartographic map, map projections make use of a so-called projection model. A projection model is formally defined as *a systematic and orderly representation of the Earth's grid upon a plane* [183]. Map projections are the mathematical mapping of the coordinates

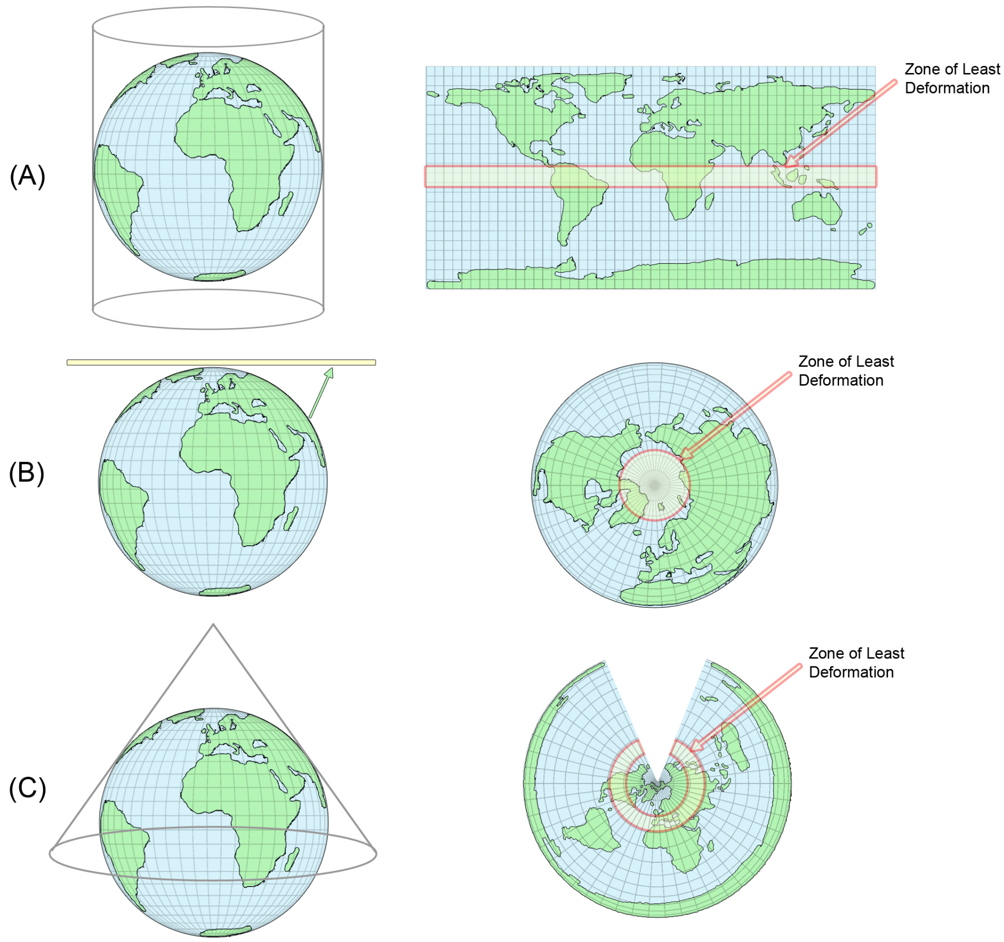


Figure 4.18: Characteristic map projections with different projection surfaces and the area of least deformation: (A) cylindrical projection, (B) azimuthal projection, (C) conic projection.

from the 3D to the 2D space, which corresponds to flattening the surface of the 3D object. Applying map projections generally involves the following three steps:

1. Choosing a suitable projection model (sphere, cylinder, plane, etc.).

2. Mapping of the geographical coordinates to a Cartesian coordinate system.

3. Scaling the map for the particular purpose.

Most of the map projections are not projections in a physical sense. They are based on mathematical formulas. For a better understanding of how map projections work, we can think of a 3D object with a light source. The surface of the 3D source object is projected by the light source onto the surface of the chosen projection model. Afterwards, the surface of the projection model is flattened by unwrapping it to a 2D surface (see Figure 4.18). A surface is called a *developable surface* if it can be flattened without distortion and without tearing the surface apart. Since not all 3D geometric shapes are developable, the choice of the projection model strongly depends on the properties that are intended for the target map. Different projection models result in different areas of least deformation. Depending on the purpose of the created map, this tremendously influences the usability of the map for that particular purpose. In general, there is an infinite number of projections possible, and more than 400 projections available (i.e., described in literature), although only a few of them are employed regularly in practice [183]. In Figure 4.18, we can see characteristic map projections with different projection models. Since there is some overlap between the different available projections, a mutually exclusive classification is not possible. Hence, map projections are commonly classified based on (1) the *preserved properties* and (2) the *projection surface.* As one effect of the fact that not every 3D shape can be flattened without introducing distortion, different map projections introduce different distortions and therefore preserve different geometric properties. Tyner categorizes the preserved properties as follows [183]:

- *Equivalence of area* (equal-area or equivalent projections). Stretching in one dimension is matched by compression in the orthogonal direction to retain an equivalent area. In this approach, angles may be distorted which leads to an altered shape of areas.

- *Preservation of angles* (conformal projections). Angles are preserved with infinitely short sides. Hence, small areas retain the correct shape and for larger areas, the overall distortion increases. To be conformal, parallels and meridians must cross at right angles and the scale has to be equal in every direction from a point. Hence, stretching in one direction must be matched by stretching in the orthogonal direction. The most prominent conformal projection is the *Mercator* projection, which is the standard map projection for nautical purposes [76].

- *Linear scale* (equidistant projections). Distances are mapped correctly. An entire map cannot be equidistant, so the distance scale of a map is correct from particular points or along lines.

- *Directions* (azimuthal or zenithal projections). Azimuths are shown correctly and the directions from a central point are preserved. Azimuthal projections usually have radial symmetry in the scales and the distortions. Measuring the azimuth between any other points is not possible.

A second approach to categorize map projections is categorization according to the projection surface. For the sake of simplicity, we describe this through the example of creating a map of the Earth (see Figure 4.18). Common projection surfaces are:

Cylindrical

When using a cylinder as the projection model, the surface of the globe is projected onto the surface of the cylinder, which is then flattened to obtain the map. On the resulting map, the latitude and longitude graticule of the globe results in a grid structure where the meridians of longitude are equally spread and the parallels of latitude remain parallel but are not equally spread. Due to the spherical nature of a globe, the least deformation for projecting a globe occurs around the equator. Cylindrical projections are well suited for spherical or curved objects. A prominent cylindrical projection is the Mercator projection (see Figure 4.18 (A)), which is the most common projection used for navigational maps. Hence, projects like *OpenStreetMap*[11] or *Google Maps*[12] apply this projection. The Mercator projection can be constructed as follows [113]: Assuming the Earth as a sphere, the coordinates x and y of a point in the Cartesian coordinate system can be calculated from the corresponding pair of latitude (ϕ) and longitude (λ) coordinates, with all angles measured in radians and λ_0 denoting the geographical length of the map center. This results in:

$$x = \lambda - \lambda_0, \tag{4.23}$$

$$y = \int_0^{\phi} \frac{dt}{cos(t)} \tag{4.24}$$

$$= \ln\left[\tan\left(\frac{1}{4}\pi + \frac{1}{2}\phi\right)\right] \tag{4.25}$$

$$= \sinh^{-1}(\tan(\phi)) \tag{4.26}$$

$$= \tanh^{-1}(\sin(\phi)) \tag{4.27}$$

[11] http://www.osm.org

[12] http://maps.google.com

Azimuthal

With azimuthal projections (projection on a plane), the surface of the globe is directly projected onto a plane. This projection has a radial symmetry in scales and distortions. Hence, azimuthal projections are well suited for mapping radial areas. The projection is constructed with a plane tangent to the globe, usually at one of the poles. The radial area around the osculation point is then mapped onto the plane. The deformation is minimal around the osculation point, and it increases with distance from it (see Figure 4.18 (B)). One prominent example is the *Orthographic* projection, which preserves neither areas nor angles. It shows the surface of the source object exactly as it can be seen from great distances. Hence, it is often used to create maps of celestial bodies and in particular the moon, which always shows the same side when viewed from the Earth. The Orthographic projection can be calculated as follows [113]: The coordinates x, y of a point in the Cartesian coordinate system can be calculated from the corresponding pair of latitude (ϕ) and longitude (λ) coordinates, and reference longitude λ_0 and latitude λ_0 by

$$x = \cos(\phi) * \sin(\lambda - \lambda_0), \tag{4.28}$$

$$y = \cos(\phi_0) * \sin(\phi) - \sin(\phi_0) * \cos(\phi) * \cos(\lambda - \lambda_0) \tag{4.29}$$

Conic

With conic projections, the surfaces of the source object are projected onto a surrounding cone which will be unfolded. Conic projections of a globe are created by putting a cone over the globe such that it is adjacent to a parallel. This parallel is called the *standard parallel of projection*. Around the parallel, the deformation is minimal; it increases with distance from the parallel. Conical projections are well suited for mid-latitude areas on a globe and circular paths around an object (see Figure 4.18 (C)). A prominent example of conic map projections is the *Lambert Conformal* conic projection, which is an angle preserving conic projection usually applied to create aeronautical maps for contact flights. The Lambert projection can be calculated as follows [113]: The coordinates x, y of a point in the Cartesian coordinate system can be calculated from the corresponding pair of latitude (ϕ) and longitude (λ) coordinates, and reference longitude λ_0 and latitude λ_0 by

$$x = \rho * \sin(n * (\lambda - \lambda_0)), \tag{4.30}$$

$$y = \rho_0 - \rho * \cos(n * (\lambda - \lambda_0)), \tag{4.31}$$

where

$$F = \frac{\cos(\phi_1) * \tan^n(\frac{1}{4}\pi + \frac{1}{2}\phi_1)}{n}, \tag{4.32}$$

$$n = \frac{\ln\left(\cos(\phi_1) * \sec(\phi_2)\right)}{\ln\left[\tan(\frac{1}{4}\pi + \frac{1}{2}\phi_2) * \cot(\frac{1}{4}\pi + \frac{1}{2}\phi_1)\right]}, \tag{4.33}$$

$$\rho = F * \cot^n\left(\frac{1}{4}\pi + \frac{1}{2}\phi\right), \tag{4.34}$$

$$\rho_0 = F * \cot^n\left(\frac{1}{4}\pi + \frac{1}{2}\phi_0\right) \tag{4.35}$$

Polygons

Polygon projections are not classical geographic map projections. Their main field of use is the *UV mapping* of texture images on 3D objects. UV mapping denotes the process of creating a 2D image representation of a 3D object in the process of 3D modeling, where u and v denote the axes in the 2D space, since x, y and z are used to denote the axes of the 3D object. The 2D texture image is mapped with UV coordinates onto the surface of the 3D object. The textured surface of the 3D object is usually cut along a manually defined seam. This seam defines which connected parts of the texture remain connected and which will be cut. This approach can be easily applied to cuboid shapes to create what we will refer to as *cubic projection*, where the 3D surface of a cube is mapped onto a 2D image without distortions. A cubic projection can in general be considered as flattening a cube by simply unfolding its surface like a box (see Figure 4.20). When projecting the surface of cuboid source objects, the area of least deformation is represented by the whole outer faces of the cube, which means that there is no deformation at all introduced (see Figure 4.19 (A),(B)). Following the example of creating a map of the Earth as a sphere, a UV mapping for a 3D model of the Earth can be constructed as follows [57]: The coordinates x and y of a point in the Cartesian coordinate system can be calculated from the corresponding x, y and z coordinates of the 3D space as

$$u = \frac{x}{\sqrt{x^2 + y^2 + z^2}} \tag{4.36}$$

$$v = \frac{y}{\sqrt{x^2 + y^2 + z^2}} \tag{4.37}$$

For each of the described categories, there is a huge number of available adapted approaches, as well as approaches that address particular properties of visualization to optimize the created maps for such dedicated purposes as visualizing large areas like the Earth. Since our goal is to create map representations of relatively small areas of various shapes, we focus on the general algorithms. In general, a map projection can be applied to a virtual model of a media facade as listed in Algorithm 3. In summary, we can say there is a huge variety of map projections available. Since they offer different features and have different characteristics, choosing a projection strongly depends on (1) the surface of the object that is intended to be mapped and (2) the purpose of the obtained map. Due to their ability to preserve different geometric properties, map projections are also highly suitable for creating 2D map representations of the surface of arbitrary non-spherical 3D objects, like buildings that are equipped with media facades.

Algorithm 3: Creating a 2D map representation of a media facade by applying a map projection.

Data: $3D$ model of the media facade
Result: $2D$ map representation of the media facade

foreach *point(x,y,z) of the media facade's surface* **do**
    x = projectToX(x,y,z);
    y = projectToY(x,y,z);
    add (x, y) to 2D map ; `// set the pixel in the 2D image`
end
return *2D map;*

4.3.3 Map Representation of Media Facades

Due to the variety of map projections that are available and due to their differing characteristics in preserved geometric properties and introduced deformations, there is no general rule for choosing the right projection for creating a map of a media facade. If the media facade covers the whole outer shell of the underlying building, the building's whole surface needs to be mapped. If the media facade covers only parts of the building's surface, we can omit the parts that do not host media facade elements from the mapping. In this case, a map projection can be chosen that best fits the characteristic geometric properties of the media facade. Different projections work with differently shaped surfaces. Media facades are not necessarily rectangular or of a distinct shape. They are integrated into the structure of a building that serves as a host. This leads to various, irregular form factors. In many cases the shape of the media facade includes different characteristic elements for which different map projection approaches are suitable (see Figure 3.2). The map projection needs to be chosen according to the geometric properties which are to be preserved. Since the goal is to obtain a 2D map of the media facade's surface that enables users to intuitively orient themselves when interacting with the facade, the main shapes and the main layout of the media facade need to be preserved. In addition to shape and layout, when interacting over the edge and leaving the part of the facade that is visible from the user's current point of view (see Figure 4.17), the borders of the areas adjacent to the ones within the current field of view of the user should also be coherent and preserved in the map representation to enable a continuous flow of interaction.

We investigated permanent media facade installations that are listed in the literature [81, 82, 26, 32, 127, 163, 103] and for which information is available throughout the web. The media facade installations are distributed around the globe. We compared their form factors, capabilities and geometrical properties to identify form factor categories that describe basic conditions which are relevant for map projections. For each category, we examined suitable map projections such that a feasible trade-off of basic shapes, the layout of the media facade and adjacent areas and edges are preserved to obtain a coherent map representation with minimal deformation. We derived the following dominant form factors of media facades for which we provide a set of guidelines to derive a facade map, and for which examples are depicted in Figure 1.3:

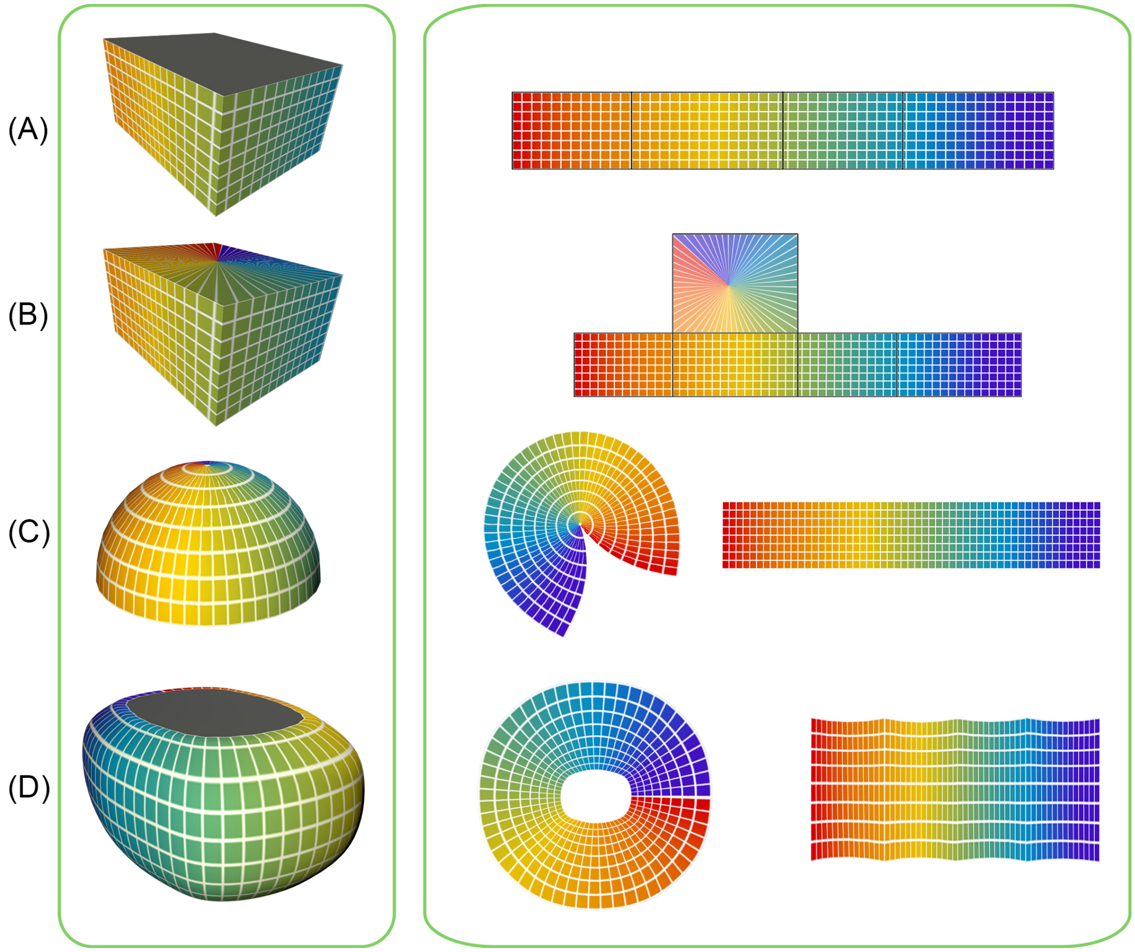


Figure 4.19: Map projections for different building shapes. (A),(B) Cubic form factor: cubic projection. (C) Spherical form factor: (left) conic projection, (right) cylindrical projection. (D) Curved form factor: (left) azimuthal, (right) cylindrical projection

Cubic

Cuboid form factors are the dominant form factors among existing media facade installations. This includes rectangular media facade installations that cover the outer shell of a building. When they cover more than one side of a building, their form creates a 3D shape that is similar to the outer surface of a cuboid. If the media facade spans around the outer shell of a building and has a cubic form factor, we have to distinguish two general cases: (1) The media facade is only formed by the outer shell (or a part of it) along the side of the building. This case can be compared to a cube without top and bottom. (2) The media facade is formed by the outer shell covering the building, including the roof.

In both cases, a well-suited map projection to obtain a 2D map of the media facade's surface would be a cubic projection, since for cuboid or rectangular shapes, there is little or no deformation introduced by the projection. If the top of the building is not a part of the media facade (e.g., the Bayer Media Sculpture[13]), the map obtained by the cubic projection is already sufficient, since the particular areas, the layout and the borders are preserved (see Figure 4.19).

In the case that the top of the building is also part of the media facade (e.g., the National Aquatic Center[14] in Beijing, China), a cubic projection with a plain unfolding of the projection shape is not sufficient since when unfolding the cube, the top side is only connected to one side part of the cube. As a result, only this one border between the top and side parts of the facade is preserved. This leads to gaps between adjacent parts of the media facade in the obtained map. If we aim for a continuous interaction with all parts of the facade, this hinders the interaction, since when the focus of interaction moves from the side parts of the facade to the top part, a continuous transition is only possible from the side part for which the border to the top part is preserved on the 2D map (see Figure 4.19 (B)). Interaction over unpreserved edges introduces a distortion in the flow of interaction. As depicted in Figure 4.20, this becomes clear when considering drawing on the facade by touch input on the 2D map. If we try to draw a circle around a corner where three sides meet, the contour of the circle overlaps with a part of the map that is not part of the media facade's surface. Since the unused parts are literally removed during the texturing, the circle that was drawn appears distorted in the direction of the border which was not preserved during mapping.

One possibility to counteract this problem could be to adjust the seam along which the surface is unwrapped to obtain the facade map, to create a *set of facade maps* where the different borders are preserved. In this case, the facade map could be dynamically exchanged while interacting, such that when the user is interacting with a particular area, all borders with neighboring areas are preserved. We believe that this might also be a drawback. Since the user directly interacts with the facade map, his interaction might be interrupted by changing the map and he might temporarily lose his orientation on the map.

[13]goo.gl/6pb6C
[14]http://www.water-cube.com/en/

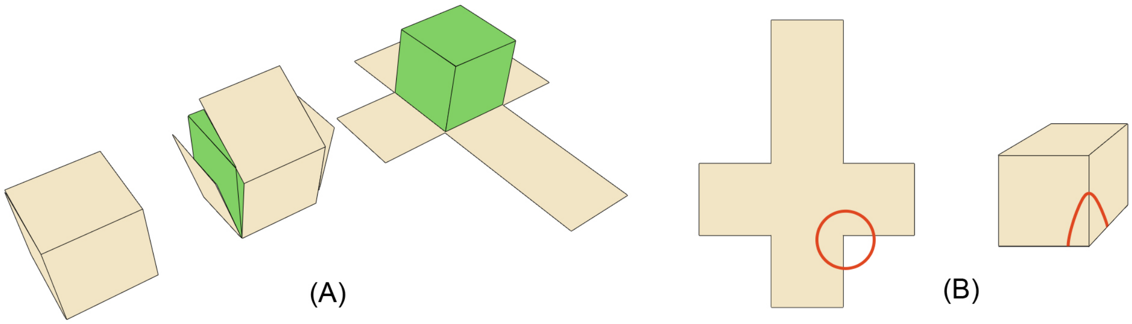


Figure 4.20: (A) Unfolding the surface of a cube to a 2D image. (B) Distortion that is introduced by interaction over unpreserved edges on the 2D image.

Spherical

We denote media facade installations that are dominated by a spherical form like a dome (e.g., the media facade of the Grand Lisboa Hotel[15] in Macau) as having a spherical form factor. For media facades with a spherical form factor, we propose applying either a conic or a Mercator projection (see Figure 4.19), dependent on the purpose of the map. If people stand in front of a spherical media facade, they usually see the lower parts of the facade from a perspective that is close to orthogonal. These parts of the facade are the primary target areas for interaction. Moving further towards the roof of the dome, the visibility decreases for a user that is standing on the ground in front of the facade. Hence, the area of least deformation of the facade map should be around the lower parts of the facade.

For conic projections, the area of least deformation is around the standard parallel. Depending on the size and the degree of curvature of the media facade, the area of least deformation tends to be towards the lower parts of the facade. The obtained map has a radial layout and is therefore well suited for vertical interactions over the top of the sphere, since the coherent structure of the top of the sphere is visible on the map. In contrast, a standard Mercator projection creates a rectangular map on which the mapped content is oriented horizontally on the map. In this case, the area of least deformation is located around the equator of the sphere. Since spherical media facades usually are only half-spheres, the least deformation will occur in the lower parts of the facade. The rectangular layout of the map makes a standard Mercator projection well suited for hori-

[15] http://goo.gl/KKZtk

zontal interaction around the facade and less suited for interacting over the top, since the coherent structure of the top is not preserved by the map. If we want to create a map that preserves the coherent structure of the top of the sphere and that is therefore suitable for interacting over the top of the sphere, we could use a *traverse* Mercator projection. This is an adaption of the standard Mercator projection with a horizontal projection axis, whereas the standard Mercator projection has a vertical projection axis. The mapped content in this case is oriented vertically on the obtained map.

An azimuthal projection would not be suited for spherical media facades since for azimuthal projections, the area of least deformation is around the tangent point of the projection plane which is generally the topmost point of the sphere. This would result in the area with the least visibility being the area of best presentation on the obtained map.

Curved

We denote the form factor of a media facade installation as curved if it is dominated by spherical and elliptical shapes that do not form a sphere (e.g., the media facade of the Allianz Arena[16] in Munich, Germany or the iluma[17] building in Singapore). For media facades with a curved form factor, the choice of a suitable map projection depends on the degree of curvature of the media facade itself, as well as its placement on the outer shell of the hosting building. If the media facade forms an elliptical ring around the host building, as for example the media facade of the Allianz Arena, a conic map projection is a suitable approach. Since conic map projections have the area of best presentation around the standard parallel, the media facade can be mapped with low distortion if the projection cone is chosen such that the standard parallel lies within the area of the media facade. If the curved shape of the media facade is rather flat, an azimuthal map projection would also be suitable to create a map. In this case, the facade is depicted as a coherent structure. For more complex curved shapes, we suggest applying a cylindrical map projection. The area of best presentation is located around the center line in the case that a sphere is mapped. If the mapped surface consists of several spherical shapes, the areas that are located around the horizontal center line of the particular shapes are mapped with the least distortion.

[16]http://goo.gl/wVsXe
[17]http://www.iluma.com.sg

For media facades where the form factor is dominated either by cubic, spherical or curved shapes, a map representation of the facade's surface can be obtained as described. If the form factor is dominated by shapes of more than one of these categories, there is a trade-off among properties that are relevant for the intended interaction and that therefore have to be preserved. One possibility could be to create a map with different map projection approaches that work best with the particular category, for each affected category, and switch during the interaction.

Orientating on the map

With a facade map, a user can interact with all parts of the media facade. This also includes the parts of the facade that are occluded from the user's current point of view. To assist the user in orienting himself on the facade map, we propose applying the metaphor of a *daylight map* to the facade map. Similar to visualizing the current area of daylight on a map of the world by displaying a shaded overlay image over the area of night, we propose displaying a shaded overlay image over the facade map to slightly shade the areas of the media facade that are occluded from the user's current point of view (see Figure 4.21). As soon as the user changes his current location or orientation, the field of view and hence the visual parts of the facade change as well. In this case, the overlay is dynamically adjusted to the current field of view of the user. To determine the location of the user and his orientation towards the media facade, we can utilize the built-in GPS and accelerometer sensors as well as the built-in compass of the user's mobile device through which he interacts with the facade. Hence, we can consider them to be available. Given that we know the location of the media facade and the geometric conditions from creating the facade map, we can dynamically estimate the current field of view of the user and automatically adjust the overlay on the facade map.

By combining the introduced concepts of map projection and texturing with the proposed guidelines for creating map representations of media facades of various form factors, the presented work creates a framework that enables new ways of continuous interaction with all parts of a media facade. This is a further step towards the exploitation of the full capabilities of media facades as urban displays and digital mediums.

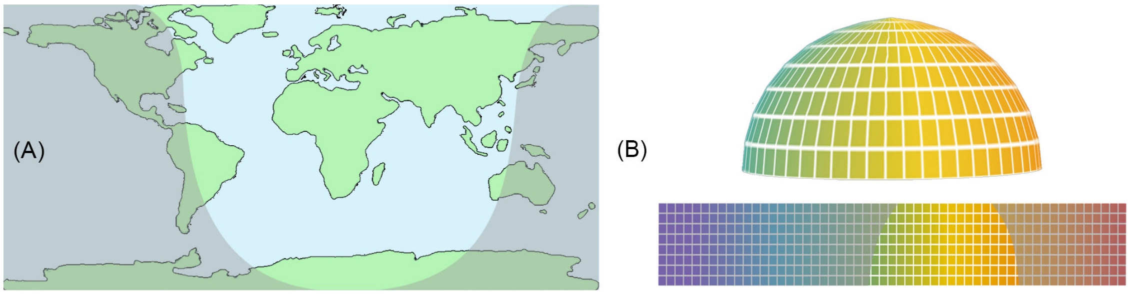
(A)  (B)

Figure 4.21: (A) A daylight map visualizing the current day and night zones. (B) The daylight map metaphor applied to a facade map to visualize the currently visible area.

4.3.4 Proof of Concept Implementation

To support the development of the proposed guidelines for creating facade maps with map projection approaches and to gain early feedback on interaction with the maps, we developed facade map prototypes and interactive 3D media facade models for the described form factors and map projections. We used the 3D modeling software Blender[18] to create 3D models of media facades with characteristic cubic, spherical and curved form factors. Since Blender offers the possibility of adding custom functionality by importing Python[19] scripts, we implemented standard map projection algorithms in Python and applied them to the 3D model within Blender to project the surface of the 3D model onto a 2D texture image. We created a UV mapping to rebind the 2D texture onto the model. To gain a realistic impression of the interaction, we implemented a client-server application for interacting with a 3D model of the facade while using a mobile phone as the input device. As a use case, we chose a painting application with which the user can freely paint on the facade. A facade map is displayed on a mobile phone and the user can paint on the 3D model of the facade by painting on the facade map on the mobile phone by direct touch input. We used the jMonkey[20] engine for OpenGL[21] to write a server application in Java that displays the respective textured 3D model of the media facade. The client application for the mobile device was written for the Android platform. The application displays the facade map and offers the possibility of painting freely on the facade with different colors and brushes that can be chosen locally through

[18] http://www.blender.org
[19] http://www.python.org/
[20] http://jmonkeyengine.com/
[21] http://www.opengl.org/

an additional menu on the mobile device. Client and server applications communicate over a wireless network connection. The client application on the mobile phone sends the necessary data, like the input events, to the server application which maps the user's input from the 2D space of the mobile phone into the 3D space of the facade model. The application was designed to support importing new 3D models with the respective facade maps to experiment with various form factors.

4.3.5 User Feedback

When designing interaction for media facades, getting early feedback and incorporating the feedback in the ongoing design process is of importance to create a usable and enjoyable outcome. Following the iterative, user-centered design process proposed by Wiethoff [192], we utilized our prototype implementation to gather initial informal feedback on general usability and the user experience when interacting with facade maps throughout the individual stages of the overall design process. During the development of the map projections for different form factors, we let our prototype for displaying the 3D model of the media facade run on a standard desktop computer, and utilized Android-based tablet computers and mobile phones on which the facade map was displayed for interacting with the 3D model. We asked colleagues and students who visited our lab to play around with the prototype. Concerning the tasks for this initial testing of the current prototype, we asked the users to mainly focus on drawing shapes and lines, and in particular to paint over edges and on occluded parts of the 3D model. Hence, possible distortions became apparent and we could get insights into how the users oriented themselves on the facade maps when interacting with occluded parts of the 3D model. Before they started interacting for the first time, we only told them that they could choose a color and a brush size and that they should paint on the facade of the 3D model by painting on the facade map displayed on the mobile device with their fingers, by touch input. Afterwards, we discussed with them in unstructured interviews what they thought was good and bad about the general approach and the particular maps that they used. We asked about how easy it was to orient themselves on the facade map as well as how they experienced the drawing (e.g., did distortion occur or not). The majority of participants stated to find distortions more acceptable when occurring towards the edges of the facade maps. When choosing between map projections, the participants also preferred the coherent part of the facade map to be as large as possible. Since this initial evaluation phase covered the whole development process, the participants got to interact with different map projections — the current prototypes — having different distortion. They provided feedback on the distortion of the particular map projections they used, in terms of how the distortion influenced drawing coherent geometric shapes. We incorporated this feedback into the further design process to refine the set of map projections proposed.

For our final set of map projections, we gathered further qualitative user feedback. We asked 10 people who own a mobile device and are familiar with using it to paint on the 3D models of media facades with the facade map prototype. The 3D model of the facade was displayed on a $15m^2$ projection wall, and the client application showing the facade map was running on a mobile device (see Figure 4.22). All participants interacted with all media facade - map projection pairs, which are depicted in Figure 4.19. We further applied the world-in-miniature metaphor [176, 199] such that every participant additionally had to perform the painting task for each facade shape directly on a miniature representation of the facade's 3D model which was displayed on the mobile device. The participants were asked to perform three main tasks:

1. In the first task, the participants were asked to paint basic shapes like circles and rectangles onto the 3D model.
2. In the second task, the participants were asked to paint onto the parts of the 3D model not visible on the screen (the backside of the 3D model).
3. In the third task, we asked the participants to continuously paint around the 3D model, and for the facade maps where unpreserved edges exist, to explicitly paint over those edges.

After completing the tasks, we asked the participants in unstructured interviews about the design and usability of our prototype as well as their experiences using it. We focused our questions in particular on qualitative aspects of the map layouts and associated issues like orienting themselves on the map and the deformation of the displayed content, as well as the connection between enjoyment and ease of use when using the prototype. In particular, we asked *how* the participants oriented themselves on the facade maps, how they experienced drawing continuous shapes and how the participants experienced the accessibility of the backside — the occluded parts — of the 3D model. Furthermore, we asked about the general usability and user experience, comparing facade maps with the world-in-miniature approach.

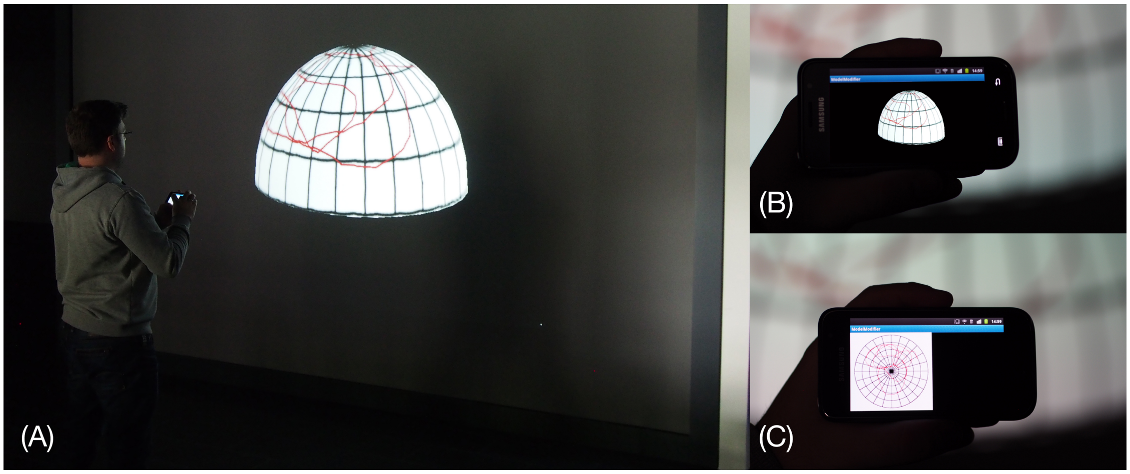


Figure 4.22: (A) A participant interacting with the 3D model. (B) The client application, showing a world-in-miniature model of the facade. (C) The client application, showing the facade map.

The majority of participants (8 out of 10) chose the facade maps over the world-in-miniature representation of the media facade as the preferred interaction technique. As common reasons for their choice, they stated that in contrast to the world-in-miniature representation, the facade map does not involve scaling and scrolling while trying to paint and they found it easier to continuously paint lines around the facade on the map. In this context, the participants also positively mentioned that the facade map helped them to get an impression on the overall content of the facade. Concerning orienting themselves on the facade map, the participants appreciated the visualization of their current field of view, by applying the metaphor of daylight maps to the facade map. When given a map where parts of the facade were displayed upside down, some participants also suggested automatically rotating the map in relation to the current focus of interaction to make sure that the current target area is always aligned horizontally. Furthermore, they requested the functionality of manually switching between different map projections. In terms of the deformation of the content when painting over unpreserved edges, the majority of the participants did not consider this as a bothersome issue. Nevertheless, the participants preferred map representations that contained as few unpreserved edges as possible. In summary, we can say that the feedback was positive throughout.

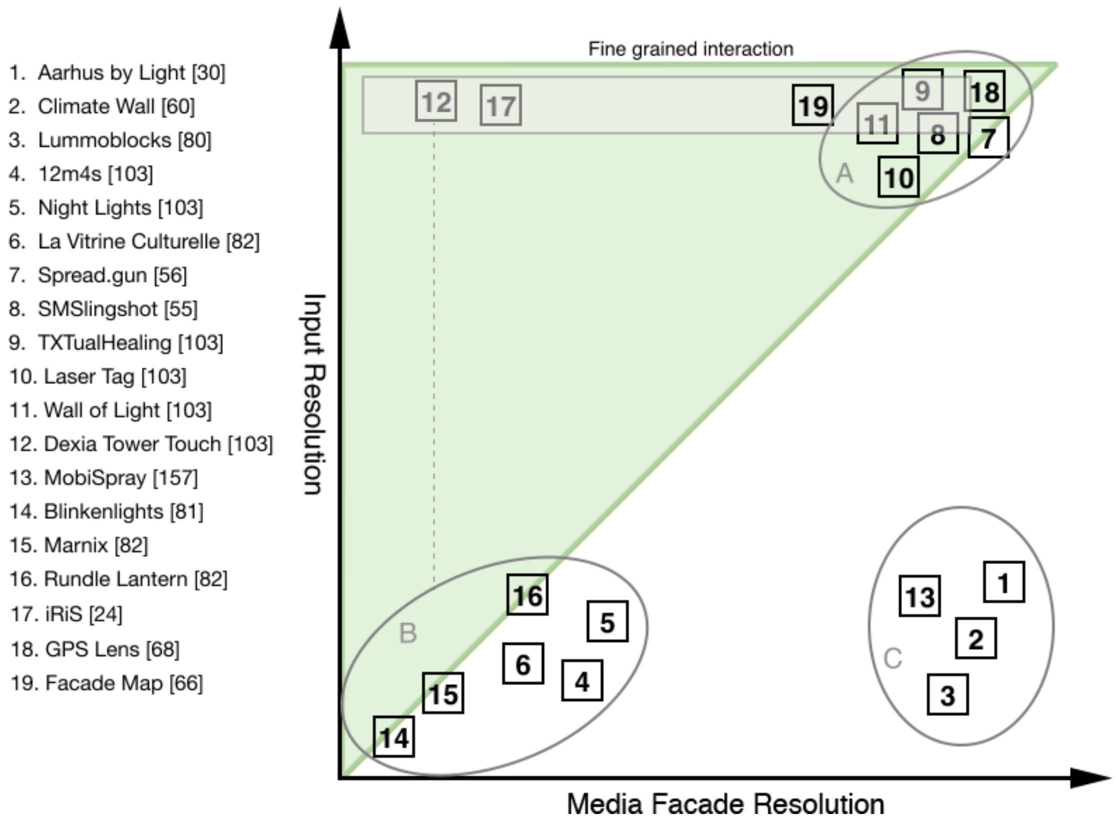


Figure 4.23: Placement of the introduced interaction technique (Facade Maps) in the classification (see also Section 3.2.1) according to input and media facade resolution (19). It is applicable to media facades of various resolutions.

4.3.6 Summary & Classification

In this section, we transferred the concepts of cartographic map projections that are used to create 2D maps to the domain of media facades. We described how to use facade maps on a mobile phone to interact with a media facade with the high precision of touch input on the mobile device's screen. We allow for fine-grained interaction with media facades of arbitrary resolutions. Figure 4.15 depicts the placement of the Facade Maps approach in the classification presented in Section 3.2.1. Furthermore, we showed how using a concept such as facade maps can help to take a further step towards exploiting the full capabilities of media facades of various form factors by enabling a continuous interaction with all parts — even the parts that are occluded from the user's current point of view — of the facade, independent of the user's current field of view. We analyzed the form factors and geometric properties of available permanent media facade installations and identified their dominant shapes. For each category, we provided guidelines for creating facade maps and proposed suitable map projection approaches. Concerning the general properties of the setup necessary to apply the facade maps concept, we need to provide a suitable 2D map representation of the media facade's overall surface. As an input device, we utilize the users' standard mobile devices. The communication between the mobile input device and the media facade is again carried out over a standard Wifi connection. Hence, we do not need any further instrumentation of the environment around the facade. This results in a placement in the taxonomy of media facade interaction (refer to Section 3.2.4) as depicted in Figure 4.24.

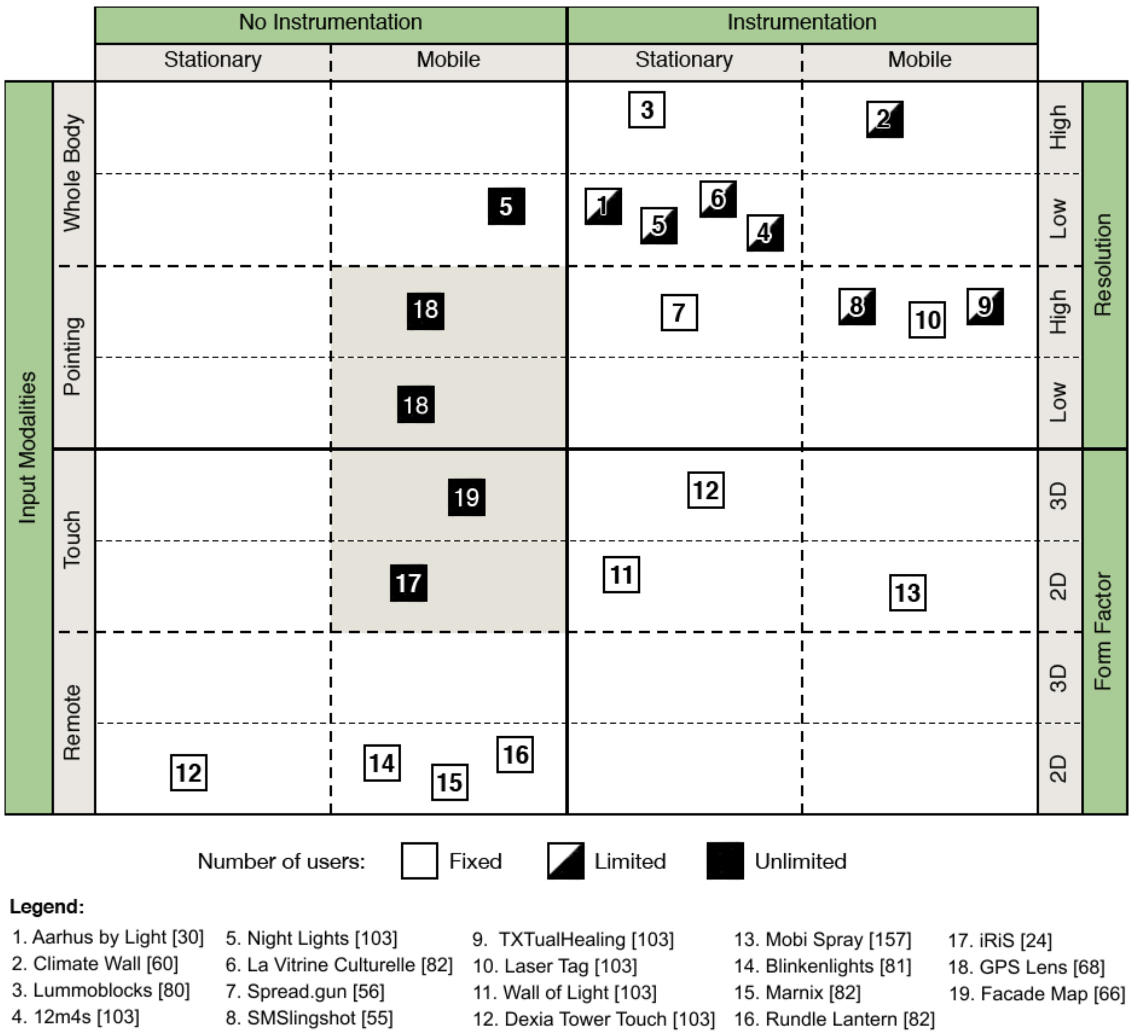


Figure 4.24: Placement of the introduced interaction technique (Facade Maps) in the classification (see also Section 3.2.4) according the characteristics and the technical specifications of the installation (19).

5

Prototyping Interaction for Media Facades

Besides the technical aspects of media facades, the context in which they are deployed and the exposure of their content to a large audience raise the need for a tailored design in terms of design and prototyping tools. Up to now, there has been a lack of development and prototyping tools for media facades. Due to the characteristics of media facades, existing toolkits are tailored to one specific setting and to the physical and technical properties of the particular media facade. They provide a simplified representation of the facade with one fixed resolution, its surroundings and the actual hardware utilized in the setting. Their main focus lies on prototyping the visual content for a media facade. Such tools usually only provide very limited functionality to prototype interaction with the media facade. If they support interactivity at all, the interaction is very basic and it usually neither supports different input modalities or interaction techniques, nor adding spatial components to the interaction by exploiting the space around the media facade. Besides inhibiting the portability of the developed installations from one media facade to another, this increases the effort needed for developing and prototyping interactive installations. Prototypes need to be fully implemented even in early stages of the design process to be used with a simulation tool, which means that they have to be tailored to the interfaces and the technical properties of the target facade. Dalsgaard and Halskov were among the first to analyze the design of interactive installations for media facades in a structured way. They identified eight key challenges for designing urban media facades [45]. They argue that the urban settings media facades are usually situated in call for new forms of interfaces or alternative assemblies of existing ones. In addition, new installations of media facades need to be integrated into the existing physical structures and surroundings, leading to complex form factors. Furthermore, the developed content needs to suit the medium, meaning the content not only needs to be appropriate to be shown in public, but also has to fit the format of the display and the kinds of interaction intended to be supported. Additionally, changes in the media facade's dynamic environment need to be considered, such as shifting light and weather conditions. Identifying and overcoming such design challenges is an integral part of the development process. Existing tools for simulating and prototyping interactive systems for urban public environments do not sufficiently support designers and developers facing the aforementioned challenges. Existing tools are mostly tailored to one specific media facade in one specific setting. They neither provide generalized means to display applications on the facade, nor do they offer possibilities to incorporate interactivity with external input devices or ways to simulate changing conditions of the area surrounding the media facade.

In this chapter we present a generalized media facade toolkit for rapid prototyping which is capable of simulating large-scale media facade installations. The simulation is independent of the size, form factor, technology and hardware of the simulated facade. We follow a modularized design approach by strictly separating the model of the underlying building, media facade, application and user interaction. With this and by allowing communication only in one way between two particular connected modules, we avoid dependencies and allow for an easy exchange of components. The underlying building model, media facades, applications running on the facades, and input devices can be individually exchanged without modifying other components. Applications can be easily transferred from one media facade to another. Within this scope, the toolkit provides means for automatically mapping applications to media facades with different form factors. The toolkit is in general capable of running arbitrary interactive applications on arbitrary media facades. It can simulate different hardware, resolutions and also different color models (i.e., additive color models for projected media facades). We describe how the modular design of the toolkit ensures application portability between different media facades and the possibility of providing interactivity by enabling user input with different modalities and different interaction devices at the same time allowing for different input resolutions (see Section 3.2.1). With the toolkit, we provide a flexible simulation, test and development tool for designers and researchers alike, which addresses the diversity of media facade settings and incorporates the media facade's architectural surroundings and the deployment space. The contributions of this chapter have been partially published in [194, 18, 65, 90].

5.1 Initial Design Explorations

We approached the design process for developing interactive installations for media facades by first carrying out design explorations where we investigated how people can get engaged to interact with a media facades. To do so, we developed a set of four prototype applications (see Figure 5.1) consisting of adaptions of well-known games such as Tetris[1] and Game of Life[2], a maze game, as well as an application to raise awareness of digital light pollution [18]:

Tetris

We adapted the original game by allowing people to use their mobile devices to control the game in two ways: (1) using buttons displayed on the mobile device's touchscreen or (2) moving around in front of the facade carrying the mobile device. With this, we introduced a performative component to the game, similar to the Lummoblocks installation [103]. The location of a user is determined with the built-in GPS sensor of the mobile device. Positions are mapped to *left*, *middle*, and *right* which are positions on the facade where blocks can be moved to. To allow for multiple users in parallel, we allowed the users to *vote* for the direction of the block.

Game of Life

With this game, we adapted Conway's *Game of Life*. A user can control a species of extraterrestrial life trying to conquer a planet (the media facade's screen) by choosing a spaceship on their mobile devices, which serves as a local launching pad for sending new units onto the facade. To do so, they point their mobile device at the facade and perform a swipe gesture on the mobile device's touch screen. The height at which the units appear on the facade is determined by the orientation of the phone which is determined with data from its built-in motion sensors. The basic game follows the rules of the original *Game of*

[1] http://www.tetris.com
[2] http://www.bitstorm.org/gameoflife/

Life. We extended the game to allow for multiple users playing simultaneously as follows: Conflicts between the different species are decided according to the *rock-paper-scissors* game (e.g., each species triangularly beats another).

Art Maze

Art Maze is a typical maze game where a maze is displayed on the facade. Users have to navigate a token from a starting position through the maze by using their mobile device as a control pad displaying keys for navigating. We allow for multiple users by displaying one distinct token per user. While moving the token around, obstacles can block the way and trap the token. In this case, the user has to perform a local task (e.g., discovering a combination of touch gestures on their mobile device) to free the token. Player-specific information — such as feedback on the gestures to be discovered in the local task — is displayed only locally on the mobile device of the particular user.

Visualization of Light Pollution

This application was developed as part of a campaign to raise the awareness of light pollution in cities by pointing out that many insects die every night because of lamps emitting UV radiation. To visualize this problem on the media facade, we display a scene of a night sky with various insects buzzing around. Using their mobile devices as local input devices, multiple users in parallel can choose between three different light sources of arbitrary size, differing in the amount and strength of emitted UV radiation: (1) high-pressure sodium, (2) low-pressure sodium and (3) high-pressure mercury. Depending on the lights chosen, insects start to *fly* from the media facade to the users' mobile devices. After several seconds, the virtual insects will vaporize as a metaphor of their death caused by that unnatural source of light. In contrast to the previous games, this installation also tries to engage passersby to interact with the media facade by conveying an educational message.

All of the aforementioned games allow basic interaction with a media facade utilizing mobile devices as input devices [18]. We explored the process of implementing different approaches to create playful engagement with the low-resolution — 192 by 219 pixels — media facade of thd PSD Bank[3] in Münster, Germany. During the design and develop-

[3]http://www.psd-medienfassade.de/medienfassade/

ment process, we experienced several difficulties in terms of (1) content development, (2) testing the applications and (3) prototyping interaction, caused by the dynamic public context of the media facade. Since the facade is owned by the company — a bank — situated in the building, all content decisions had to be continuously coordinated with the advisory board of the company throughout the development process. Furthermore, the final applications with respect to content, visualization, stability and interaction also had to be approved before each deployment. For this reason we were asked to create realistic prototypes which can be presented without running the application on the actual facade. As for testing application and interaction before the deployment, this turned out to be a challenging task. Since all content deployed onto a media facade is already exposed to a wide audience and since testing time is also highly limited due to the frequent occupation of the facade, we had to reproduce a testing environment in a lab setting. The challenge here is to reproduce the setting as accurately *as possible* and at least as accurately *as necessary* for testing particular parts of the installation. The necessary accuracy in this case is determined by the parts of the installation that are to be tested. These parts must be modeled such that the testing can provide reliable results.

Due to their size, the technical realization and the highly dynamic environment around media facades, it is not possible to reproduce an identical test setting. Hence, we developed a virtual simulation tool for this particular facade, modeling the parts of the setting that are necessary to simulate how the application will look on the actual facade and how users can interact with it utilizing their mobile devices as input devices. The simulation showed a virtual 3D representation of the media facade — including the host building — on which the developed applications were running. While this approach worked well

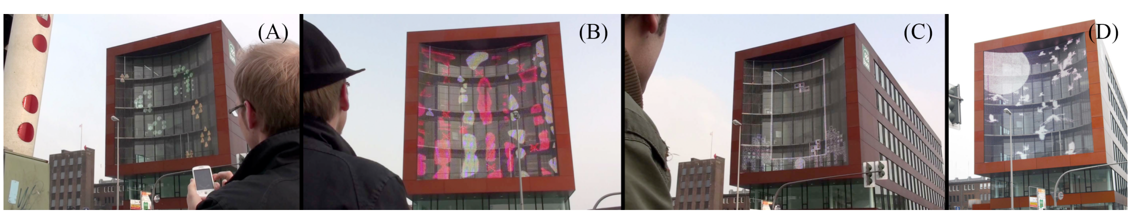


Figure 5.1: Four design explorations for basic interaction with a media facade: (A) *Recolonize*, an adaption of Conway's Game of Life, (B) *ArtMaze*, a maze game with quests, (C) a *Tetris* clone and (D) an interactive visualization for a campaign against light pollution.

to test the overall functionality of the application, as well as its visual appearance (especially aspect ratio and layout), it was not suitable for providing an impression of the visibility of the content from different locations around the facade, as well as the interaction from those places. Furthermore, the simulation was not capable of simulating different conditions of the environment — such as changes in lighting conditions — which influence the usability and experience of the installation in the final deployment setting. This turned out to be especially important when designing installations for light emitting media facades, since lighting conditions heavily influence the visibility the displayed content.

Project iRiS

During the development of iRiS (see Section 4.1), our installation for the ARS Electronica Festival, we approached the development following a user-centered design approach using purpose-built prototyping toolkits including both hardware toolkits and software tools, and we repeatedly gathered user feedback during the particular stages in the design process [194]. To investigate the users' needs and their experiences while interacting with the final application on the media facade, as well as with the particular prototypes in the different development stages, we made use of and adapted existing *user experience* (UX) evaluation methods. In [192], Wiethoff combines the applied approaches we described in [194] to derive design and evaluation methods. Figure 5.2 depicts the adapted user-centered design process [130, 155] described by Wiethoff. This design process focuses on extending the experience prototyping and evaluation phase (highlighted in green) to suit the context. It allows to (1) test and explore the interaction concept and underlying hardware before deploying it on the actual facade, thus performing more design iterations [35], and (2) make experimental use of user experience (UX) [8] in combination with human computer interaction (HCI) [104] evaluation methods to investigate the implemented interaction concept holistically. To follow, we will focus on the particular prototyping tools and our experiences using them.

As pointed out previously, compared to traditional graphical user interfaces (GUI), the design of interaction for media facades comes with additional challenges. In particular, the physical properties such as size of the display, resolution and the used display technologies, but also testing and exploring both concepts and hardware before the final deployment, were challenges we had to face during the development process. With the facade being operable only a few hours per day and being frequently occupied, we had to align our prototyping efforts in such a way that we did not need access to the facade itself. Independent of designing the content to be displayed itself, the first challenge that has to be faced in the design process is how to access the media facade and how to control it. For prototyping, we have to model the technical setup and the hardware interplay as accurately as possible.

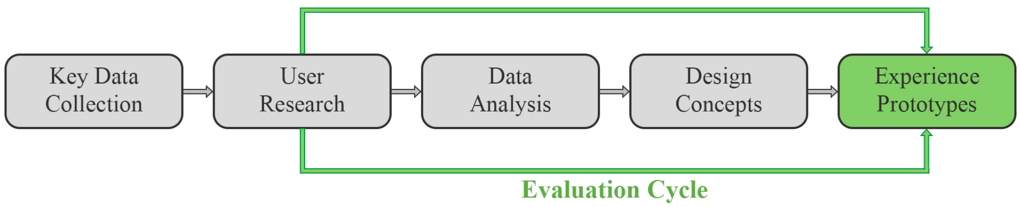


Figure 5.2: The extended user-centered design process according to Wiethoff [194, 192]. The design process focuses on extending the experience prototyping and evaluation phase (highlighted in blue) to suit the context of media facades.

Wiethoff and Blöckner introduced Lightbox [193], a hardware toolkit aiming to provide designers a way to test the hardware interplay of the particular technologies used to assemble a media facade, as well as the input devices enabling interaction with the facade. Lightbox consists of an aluminum box measuring 48 x 38 x 25cm. The lid of the box holds a panel of 12 x 12 LEDs, built with 12 single 24V / 10W high-power colormix RGB LED strips (see Figure 5.3). The LED strips are controlled through DMX signals, an industry standard for controlling lights, which is also used at our target facade. Additionally, Lightbox is equipped with a computer running dedicated software to control the box, as well as a 24V and a 9V power supply to power and control the LED panel and the experimental setups. We applied Lightbox to prototype the hardware interplay of our installation as it closely simulates the low-resolution facade of the ARS Electronica Center. We used the same hardware elements and communication protocols used for the actual facade in order to ensure our prototype provides full compatibility with the deployment environment. Concerning the visual appearance of the installation, Lightbox could only be used to get a general impression of how an application might look on such a low resolution facade and to test the interaction before deployment onto the facade itself. However, its main purpose is simulating hardware and its visual prototyping capabilities are limited. Furthermore, it covers only a small portion of the full media facade and it also has a different resolution.

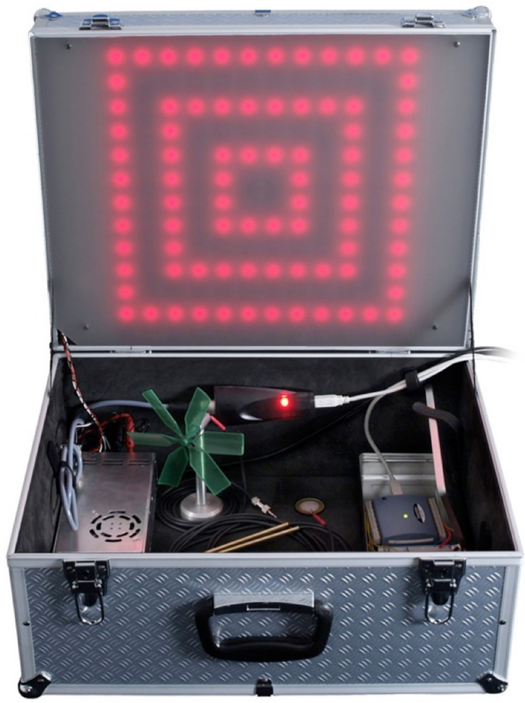


Figure 5.3: Pre-testing design concepts using Lightbox [193], a mobile prototyping toolkit to simulate the hardware interplay of interactive media facade installations.

To prototype the visual appearance of an application before the deployment, ARS Electronica provides a software tool allowing to explore content on a 3D rendered scene of the building and its static environment (see Figure 5.4), not including aspects like different points of view, interactivity, or changes in lighting and weather conditions. However, simulating the application on a liquid cristal display such as a regular computer screen was different from testing on the actual facade. The brightness levels and especially the visibility within the space around the facade did not give a realistic impression of the behavior in the real setting. For example, emotional experiences triggered by colored, bright lights are completely lacking. The tool did not account for light sources in the simulation environment, which could give an impression on how light falling onto the facade and being reflected on it can influence the visibility of displayed content and different colors. According to Wiethoff [192], they are important factors, since when interacting with a media facade, people usually tend to prefer enjoyment over the ease of use of the installation [194]. Furthermore, the lack of possibilities for simulating changes in lighting and weather conditions with the provided tool became apparent while testing on-site at the ARS Electronica facade. Wet ground caused by rain in front of the facade lead to reflections from the facade. Since the interaction technique we tested relied on visual means (see Section 4.1) (e.g. visual detection of a white frame around the facade), the reflections on the ground were falsely detected as the facade, which prevented the system from working.

Our efforts to overcome the aforementioned challenges with the limited capabilities of existing means for prototyping interaction with media facades, as well as the fact that the developed applications and the prototyping tools themselves are highly tailored to one specific setting, encouraged us to develop a generalized simulation environment, which we will introduce to follow. Before introducing the media facade toolkit, we outline the general limitations of current toolkits.

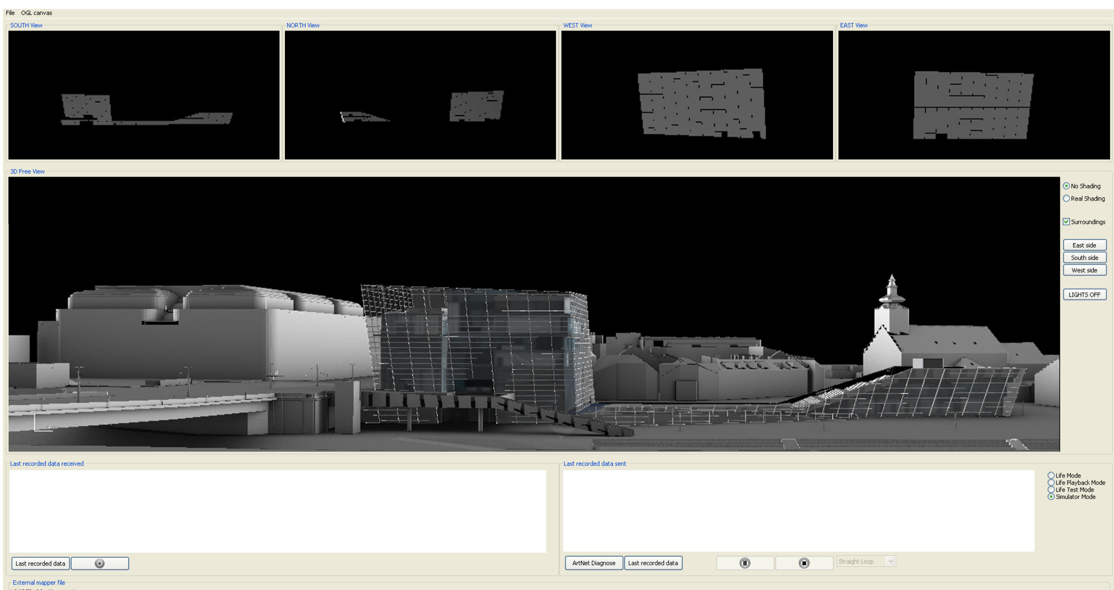


Figure 5.4: The ARS Electronica simulator to visualize applications on a 3D representation of the ARS Electronica Center's media facade (http://www.aec.at). The tool only visualizes the static architectural surroundings without lighting conditions.

5.2 Limitations of Current Toolkits

As pointed out in the previous section, when designing interaction with media facades, it is a common approach to build prototyping toolkits using simplified *small-scale* models of the media facade. This comes with further limitations. Since they cannot sufficiently map all features of a media facade, they are usually built for modeling one specific feature. As we experienced when using Lightbox [193], the toolkit could be used to prototype and test the hardware interplay of the particular components while it only offers very limited means for visual prototyping. It can only model 2D excerpts of a media facade with a limited mapping of the surrounding scenery and environmental conditions influencing the visibility of the application when interacting with it in the final setting. Furthermore, Lightbox is a custom prototyping toolkit, tailored to one specific setting. This also comes with the requirement to continuously update the prototyping tools to match the hardware setup of the facade.

A software-based toolkit could avoid many of the issues introduced by physical toolkits like Lightbox, or, when used in conjunction with hybrid prototyping toolkits like the *City Compiler* of Nakanishi et al. [135], alleviate them to a certain degree. With a modularized simulator, the prototyping environment could be easily adapted to new settings without loosing simulation precision and without binding the application to one particular setting. However, currently available software tools are at least as specifically tailored to the target facade as physical toolkits. Furthermore, they often only include an abstracted, static representation of the media facade's environment without providing options to dynamically simulate changes in the environment. However, this has turned out to be an important aspect for the design of installations for media facades. Due to their hard-coded nature and the often missing abstraction between application and facade, they force the programmer to fully implement prototypes even in the early stages of the development process in order to test them. To be able to use such a simulation tool during the development process, a developer is forced to commit his application to this specific installation, resulting in the issue that transferring the application to another facade is not possible without re-writing large parts of the application.

To overcome this, usually both low-fidelity prototypes representing simplified and abstracted models and high-fidelity prototypes mapping the target setting as accurately as possible are created from interactive systems during a design process [153]. The choice of prototypes depends on (1) how accurately they need to represent their real-world counterparts, and (2) the details and functionality to be addressed. Both are influenced by the particular stage in the design process. The choice depends on the aim designers and developers are pursuing with a prototype, which is commonly referred to as the prototype's scope [115]. Such tools help designers to prototype applications early in the design process in a more efficient and cost-effective manner and they provide a high degree of customization. When designing for media facades, however, there is no common ground for creating such prototyping tools. We need to ask ourselves why it is so difficult to build a toolkit powerful enough and flexible enough for the development and simulation of interactive installations for media facades, which can be transferred from one media facade to another — independent of the media facade's form factor and its technical specifications — without tailoring the developed application to one specific setup.

The answer lies in the huge variety of media facades, their unique characteristics and the dynamic public context in which they are deployed. While certainly not all features can be sufficiently captured within a prototyping environment, we need to identify the key features that are essential for providing a generalized simulation framework. We build upon the aforementioned ideas to provide a more general, flexible and powerful approach which supports the integration of interactivity with different modalities and input devices, as well as the full integration of existing applications into a virtual representation of a media facade. Furthermore, we want to ensure that with the media facade toolkit, developers can create applications independent from a particular media facade, such that the application can be deployed to different media facades with different capabilities and properties with ease. In Section 4.3, we introduced the idea of applying cartographic map projections to create 2D map representations of the 3D surface of a media facade [66]. We integrate this idea into the media facade toolkit to dynamically map 2D application canvases onto arbitrary-shaped media facades. By doing so and by designing the toolkit in a modular manner, such that user input, application, facade and 3D model of the building are strictly separated, we ease the portability of applications from one media facade to another and reduce dependencies between the particular components of the toolkit.

5.3 Media Facade Toolkit

When designing interaction in general, Nielsen recommends an iterative design approach and the use of low- and high-fidelity prototypes, addressing specific features of the overall system with different dedicated prototypes [136]. For regular graphical user interfaces running on standard high-resolution computer screens with a rectangular form factor, researchers can easily construct a prototype and choose from a variety of tools and approaches ranging from GUI prototyping frameworks to tangible toolkits [35]. For public environments with the complexity of media facades, this is nearly impossible due to their specifications and the highly dynamic and public context they are situated in. Due to the unique configuration of media facade installations, it is a challenging task to provide generalized prototyping tools that are not tailored to one specific facade or at least specific characteristics and features. However, when designing interactive installations for media facades, the possibilities for prototyping installations in general are also limited. Due to their technical specifications, light emitting media facades in particular are mostly not visible and active during daylight. This restricts time for testing during the development process to a few hours per day. In addition, the deployment of early prototypes for testing is already visible to a large audience since media facades are usually situated in urban public spaces. As they are usually owned by companies or public institutions, the stakeholders of a media facade further usually do not want early prototypes to be exposed to the public as they are afraid of malfunctioning prototypes. They also fear that the lack of visual appeal of early prototypes might have a negative influence on the public opinion on the company or institution. As a result, not many design iterations are feasible on the media facade itself and the particular design iterations usually happen late in the design process, resulting in high levels of effort when making changes to the media facade installation.

Hence, researchers and designers are forced to do most of the testing in an artificial lab setting, using projected or regular displays with different properties than the actual facade, which usually causes discrepancies between the prototypes and the final deployment setting. The physical and technical properties of a media facade prevent building a full-scale replica for development. Even in a smaller scale, this is nearly impossible while retaining the full set of functionality of the original facade. In summary, we have to face the following challenges during the development and deployment process of interactive installations for media facades. These have been derived from the design challenges described by Dalsgaard and Halskov [45]:

1. **Various irregular form factors**: Media facades are embedded into the physical and architectural surroundings of an existing building which results in various 3D form factors matching the surface and shape of the building. In combination with the great visibility of media facades and the associated variety of possible viewing angles often resulting in occlusion of parts of the displayed content, this raises challenges for the design of the facade itself (e.g., the pixel density and layout of the pixels), as well as for the design of content.

2. **Robustness and stability**: Depending on the utilized display technologies (e.g., LED screen elements, projections, etc.), the facade's content might not be visible under certain conditions such as direct sunlight being reflected on the facade. In this context, contrasts and displaying colors are important aspects for prototyping in order to provide robust media facade installations.

3. **Limited testing**: Light emitting media facades are usually not visible during bright daylight which limits the time-frame for testing to a few hours per day. The frequent occupation of media facades limits this time-frame for testing even further. In addition, all content displayed on the media facade is immediately exposed to a large audience. Hence, stakeholders usually only allow well-designed prototypes to be deployed onto the facade. This characterizes the late stages of the design process.

4. **Content development**: The content has to (1) fit the size, resolution and form factor of the display and (2) it needs to be appropriate in a highly public context. The variety of possible viewing angles and distances also needs to be addressed, in order to reach the the audience.

5. **Interactivity**: Since direct touch input usually is not possible, new forms of interaction at a distance need to be integrated. During the design process, we need means to prototype and test interactivity from different distances and viewing angles with a varying number of users.

6. **Portability**: With existing development tools, applications are highly tailored to one specific setting. Due to this strong connection between facade and development tools, the development tools have to keep up with changes in the facades in order to ensure the compatibility of the developed apps with the target facade. Furthermore, when transferring installations to other facades, enormous effort is required to adapt the installation accordingly.

The aforementioned challenges can be addressed by creating special-purpose prototyping toolkits, addressing some of the challenges, but not all of them. However, existing toolkits still come with certain limitations. With the media facade toolkit, we address the aforementioned challenges. In his classification of media facades, Haeusler [81] lists a number of distinguishing factors, including their size, shape, form factors and display technology, which in particular need to be considered when creating prototyping toolkits. Our toolkit needs to capture these properties as accurately as possible. Since it is the core of the visualization, the building itself into which the media facade is embedded as well as its architectural surroundings naturally also need to be captured in the simulator. The physical properties of the building, as well as the surrounding architectural space, dictate possible viewing angles and distances for a user interacting with content displayed on the facade. They further influence the general visibility of the media facade and hence the number of possible users. Furthermore, the scale, size and form factor of the media facade are also important factors for correctly mapping and displaying content on the facade. For example, for media facades with a 3D form factor, content originally designed to be displayed on a 2D canvas might appear distorted, or parts might be occluded when displayed on the media facade. As pointed out by Haeusler [81], media facades can be categorized by their technical specifications. The utilized display technology influences how an application performs under different lighting conditions. For example, colors might be perceived differently and the general visibility of the content might change under different lighting conditions. Furthermore, in order to develop interactive applications for media facades, user input and interaction with an arbitrary resolution factor Res (see Section 3.2.1) need to be adequately modeled in the simulation, and the spatial aspects of the interaction need to be included.

As pointed out in Section 3.2.4, the driving parameters for interactive media facade installations are the utilized input modality and the way the interaction is carried out (e.g., stationary or mobile), the instrumentation of the potential interactive space, the resolution of facade and user input, as well as the form factor of the media facade. To comport with this, a general prototyping and simulation toolkit for interactive media facade installations needs to support

1. any number of users using
2. any class of input device to
3. control an application
4. running in real-time on a media facade of
5. any technology, shape and size. The media facade can be
6. embedded in any building with
7. surrounding architecture.

Due to the characteristics of media facades and the dynamic nature of the environment they are deployed into, addressing all these issues to a full extent in a simulation is a challenging task. However, we can provide a toolkit that allows prototyping and simulating interactive applications on arbitrary media facades by designing the toolkit using a modularized design approach, providing a high level of flexibility. The central part of the toolkit is the simulation of the virtual setting comprising all modeled components. The simulation is an essential part of the development and prototyping process of an interactive media facade installation as it provides immediate visual feedback. However, it can be used in the same way detached from the development process to visualize existing applications in different settings.

5.3.1 Architecture

The media facade toolkit follows an extended *model, view, controller* pattern, which was introduced by Krasner and Pope [106]. We believe that exploiting such a modular design pattern is the key to providing a flexible toolkit. As can be seen in Figure 5.5, the toolkit is designed with a modular approach, strictly separating building models, media facades, applications and interaction. By establishing abstraction layers, specifying the communication format between the modules and allowing communication in only one

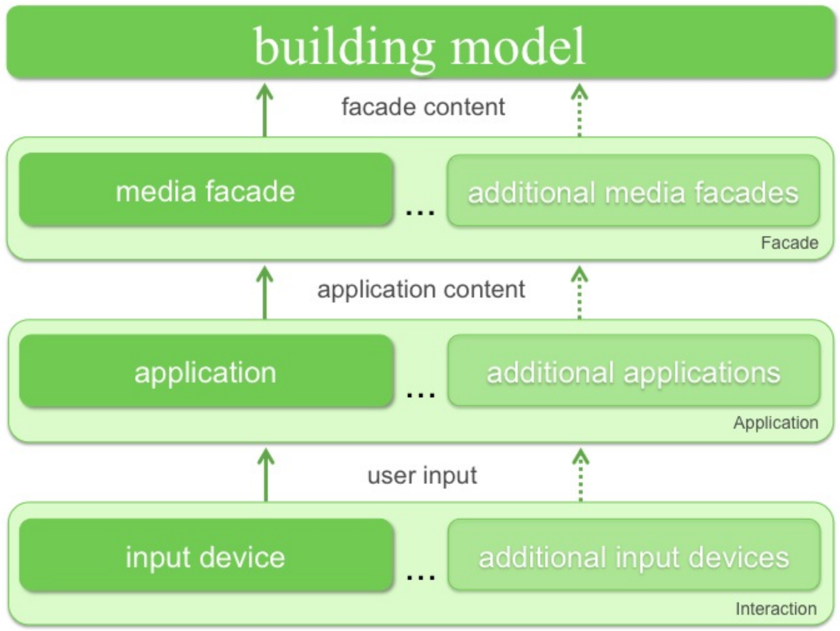


Figure 5.5: The structure of the toolkit. The building model, media facades, applications and input devices are organized in separate modules, retaining a flexible structure. Communication takes place only between pairs of modules. Input devices send user input to the applications, the applications communicate their visual content to media facades and they communicate the facade content to the building model.

direction between two particular connected modules, we organize the modules hierarchically to prevent dependencies while providing means to communicate and send data to connect the modules. The toolkit's modules communicate as follows (see Figure 5.5): Input devices send data to applications. Applications send their content to facades, and only facades communicate with the model directly. By doing this, we further provide the possibility of replacing instances of certain modules while retaining all other components. For example, the building hosting a media facade can be replaced easily on the fly and without modifying any other component, such as the application or the media facade itself. Hence, applications can be easily transferred from one media facade to another. By limiting the communication to the next element in the hierarchy, we also limit the number of abstraction layers needed and reduce dependencies between the different components of the toolkit. We now give a detailed description of the modules contained in the toolkit.

Interaction

User input is directly handled by the toolkit and not by the particular applications. With handling user input directly toolkit, we can support various interaction techniques and input devices, allowing for interaction with arbitrary input resolution on media facades with different display resolutions. The toolkit supports a certain set of toolkit commands (see Table 5.1), which offers the advantage that we only need to create a mapping from input devices to toolkit commands. As a result, specified input devices and techniques can be reused for various toolkit applications with ease. With this, we can simulate different resolution factors *Res* by utilizing different interaction techniques and devices for a given media facade. The interaction module of the simulator defines the supported interaction techniques and specifies the mapping of user input to simulator commands. As the interaction and application modules are strictly separated, a developer can develop and specify the input and interaction modalities as so-called input senders, independent from the actual application. An application only needs to specify which internal commands, signals or gestures supported by the toolkit are supported within the application. To be independent from any input device, applications can register and listen for input signals according to the *Observer Pattern* introduced by Gamma et al. [61]. This pattern defines a one-to-many dependency between a subject object and any number of observer objects so that when the subject object changes state, all its observer objects are notified and updated automatically.

```
<SimulatorInput type="Keyboard">
        <input>
                <signal>LEFT</signal>
                <code>LEFT_ARROW</code>
        </input>

        <input>
                ...
        </input>
        ...
</SimulatorInput>
```

Figure 5.6: Specification of user input. The supported input signals of the toolkit (e.g., LEFT) are mapped onto input codes of external input devices (e.g., LEFT_ARROW). An input mapping needs to be created for every supported input device.

The developer can specify in XML (http://www.w3.org/xml) notation how the signals sent from an arbitrary input device such as a keyboard, mouse, mobile device, etc. map to the supported input commands of the application. An example of a mapping from external to internal signals is depicted in Figure 5.6. Input senders are not necessarily concrete physical devices held by the user. All possible interaction techniques and devices can serve as input senders as long as their output can be encoded in the described mapping. Dalsgaard and Halskov [45] give examples of installations that also support whole body and gestural interaction without dedicated input devices. Such techniques are supported by the toolkit as well.

Buttons	BUTTON_0 , BUTTON_1 , BUTTON_2 , BUTTON_3 , BUTTON_4 , BUTTON_5 , BUTTON_6 , BUTTON_7 , BUTTON_8 , BUTTON_9 , BUTTON_10 , BUTTON_11 , BUTTON_12 , BUTTON_13 , BUTTON_14 , BUTTON_15 , BUTTON_16 , BUTTON_17 , BUTTON_18 , BUTTON_19 , BUTTON_20 ,
Number keys	NUMBER_1 , NUMBER_2 , NUMBER_3 , NUMBER_4 , NUMBER_5 , NUMBER_6 , NUMBER_7 , NUMBER_8 , NUMBER_9 , NUMBER_0
Typical special keys	ESCAPE , RETURN , LEFT_BRACKET , RIGHT_BRACKET , LEFT_ALT , RIGHT_ALT , LEFT_CONTROL , RIGHT_CONTROL , LEFT_OPTION , RIGHT_OPTION , LEFT_SHIFT , RIGHT_SHIFT , COMMA , PERIOD , SLASH , BACKSLASH , SEMICOLON , APOSTROPHE , GRAVE , SPACE , CAPSLOCK , BACKSPACE , TAB , END , PAUSE , HOME , PRIOR , NEXT , INSERT , DELETE

Letter keys	Q, W, E, R, T, Y, U, I, O, P, A, S, D, F, G, H, J, K, L, Z, X, C, V, B, N, M
Operator keys	MINUS, MULTIPLY, SUBTRACT, ADD, DIVIDE, SCROLL, EQUALS, DECIMAL
Function keys	F1, F2, F3, F4, F5, F6, F7, F8, F9, F10, F11, F12, F13, F14, F15
Numpad	NUMLOCK, NUMPAD_1, NUMPAD_2, NUMPAD_3, NUMPAD_4, NUMPAD_5, NUMPAD_6, NUMPAD_7, NUMPAD_8, NUMPAD_9, NUMPAD_0, NUMPADEQUALS, NUMPADENTER, NUMPADCOMMA
Arrow keys	UP_ARROW, DOWN_ARROW, LEFT_ARROW, RIGHT_ARROW
Touch events	TOUCH_TOUCHED, TOUCH_MOVED, TOUCH_UNTOUCHED, TOUCH_MERGED, TOUCH_SPLIT
Axis	AXIS_X, AXIS_Y, AXIS_Z, AXIS_A, AXIS_B, AXIS_C, AXIS_D, AXIS_E, AXIS_F, AXIS_G, AXIS_H

Table 5.1: The toolkit commands supported by the Media Facade Toolkit

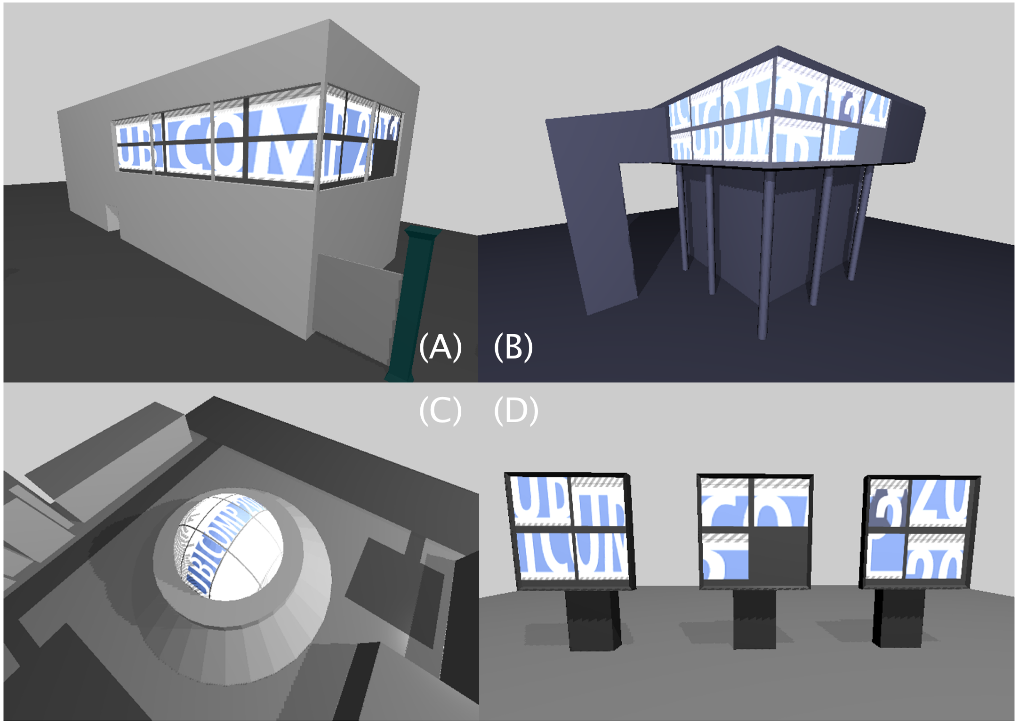


Figure 5.7: A puzzle game mapped to four media facades with different form factors by simply exchanging the underlying building model without implementing changes to the actual implementation. (A),(B): Rectangular facades. (C) A spherical facade and (D) a rectangular facade which consists of three non-coherent parts.

Application

An application can be any program producing visual application content that can be displayed on a digital screen. An application can support interaction by listening for and reacting to user input from a specified interaction module. Only toolkit commands need to be supported and there is no need for the designer of the application to design dedicated input devices and interaction techniques. The number of parallel users in this case is not restricted by the toolkit, but it may very well be restricted by the application itself, if desired. Due to the modularized design of the toolkit, an interactive application supports all interaction techniques and modalities specified in the toolkit at the same time, without forcing the developer of an application to choose one. The input commands only need to be specified once as described in Figure 5.6. However, support for interactivity is optional. The toolkit also supports visualizing videos, animations or general non-interactive visual content on a media facade. Figure 5.7 depicts an application which was mapped onto four different media facades by simply switching the underlying building's model while keeping all other components the same. As can be seen, the application, originally developed to be displayed within a 2D rectangular canvas, is displayed on a rectangular facade (A,B), a spherical facade (C), and a rectangular facade which consists of three non-coherent parts. The displayed content is automatically distorted and mapped onto the differently shaped facades. The automatic mapping is calculated using the facade maps approach described in Section 4.3. The only requirement for an application to be supported by the toolkit is that it has to extend a provided toolkit application class (see Figure 5.8). An application needs to respond to a trigger to execute one step. Furthermore, it has to provide functionality for correctly closing the app within the toolkit. The most important functionality that has to be provided is that an application needs to continuously render its visual content into a provided application frame buffer. Optionally, if interactivity should be provided, an application needs to listen for and handle toolkit commands for user input.

```
public class MY_APP extends Toolkit_SwingBasedApplication {

    @Override
    public void respond(long millisecondsSinceLastSignal) {
        // Trigger the application to execute one step
    }

    @Override
    public void render(Framebuffer framebuffer) {
        // Copy the visual content into a framebuffer
    }


    @Override
    public void quit() {
        super.quit();
        // What needs to be done to quit the application?
    }

    @Override
    public void receive(InputCode code, int value) {
        // Respond to user input
    }
}
```

Figure 5.8: Applications need to extend a toolkit application class. They need to execute application steps, render their visual content into a provided framebuffer, clean up once they are closed and optionally handle toolkit input commands.

Media Facade

Within the toolkit, a module which transforms the visual output of an application into content that can be displayed on the building model is denoted as a media facade. The visual content of the application is rendered into a separate frame buffer (see Figure 5.8). For the simulation, the properties of the facade obtained from the virtual 3D model such as its resolution, the supported color model or lighting conditions of the environment are taken into account. Although a model can include multiple media facades, one media facade can show at most one application. In a case where multiple applications are intended to run on the same media facade, the media facade can be subdivided into individual media facades, each holding one application.

Displaying the content of an application on a media facade is realized by mapping the application's content frame buffer onto the media facade in the 3D model as a texture. First introduced by Catmull [39], texturing has become a common, widely used technique. The term texturing denotes the process of covering the surface of a 3D object with 2D images which are called *textures*. Heckbert defines a texture as *a detailed pattern that is repeated many times to tile the plane, or more generally, a multidimensional image that is mapped to a multidimensional space* [93]. This allows 2D content to be displayed on a model with a 3D form factor. The textures are mapped using texture coordinates, mapping each pixel of the texture image to a point on the surface of the model. Applications developed for a regular, rectangular digital screen, such as a common desktop monitor or a situated public display, usually have a well-defined application frame with a quadratic or rectangular shape. Hence, this results in a quadratic or rectangular texture when mapping such an application onto the surface of a 3D object. Within the media facade toolkit, the application's content frame buffer generally has a well-defined rectangular shape as well. By supporting transparency, the toolkit also allows for mapping irregular shapes onto a facade. This can be used to map applications onto media facades with a non-cohesive shape (see Figure 3.1 B). In this case, the application frame buffer can be considered a rectangular bounding box for an arbitrary and not necessarily coherent form.

A user can mark every surface in the 3D model as a possible media facade by naming it as such in the model. Stahl and Haupert illustrate a way to capture an application's visuals, mirroring it into a virtual display in real-time [173]. However, the virtual displays in this case have a rectangular shape and a 2D form factor, resulting in a trivial mapping.

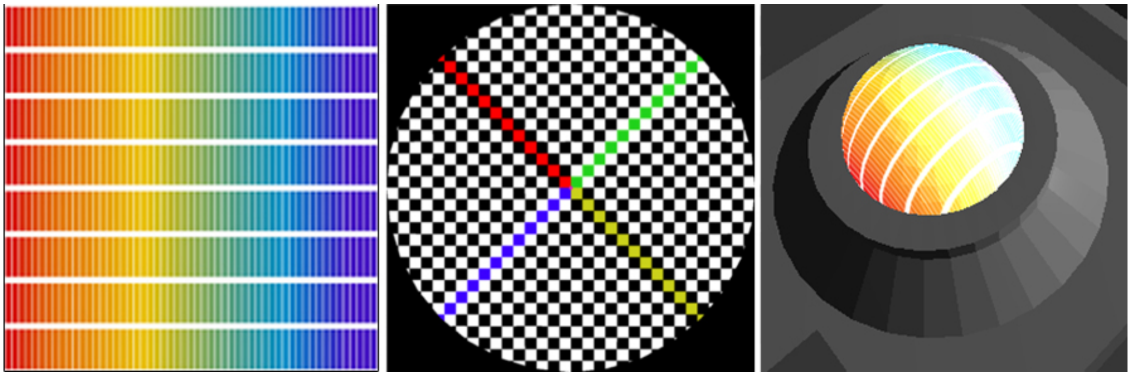


Figure 5.9: Mapping an application onto a media facade model as a texture. To get a better impression of the mapping and the introduced distortion, the texture shows a test image rather than an actual application. Left: the texture image to be mapped onto the model. Middle: The 2D image created from the 3D model surface by applying map projections. Right: The texture image mapped onto the 3D model.

To compute the texture mapping for mapping a media facade's visual content onto the building model, we use the facade maps concept (see Section 4.3) to apply cartographic map projections. When selecting a surface in the 3D model as a media facade, a map projection is applied accordingly to create a 2D image of this surface with the associated texture coordinates for the mapping. The advantage over common texturing using UV coordinates is that in contrast to UV mapping where the texture is optimized in terms of the size and distortion of the texture (e.g., the obtained texture image usually consists of several non-coherent parts), we obtain a coherent texture image which is an important aspect for the automatic mapping of a media facade onto the surface of a building. To map the facade onto the building, we invert the projections to map a 2D image onto a 3D surface (see Figure 5.9). By doing so, the facade automatically appears correctly distorted on the building's model. For any point on the face in 3D space, a matching point in 2D space can be interpolated. The texture is calculated out of a finite number of pixels but the interpolation result is continuous; therefore the color value of this point on the texture often comprises multiple pixels, e.g. bi-linearly filtered in the x- and y-directions. This value is incorporated during the shading process. The final color of a fragment is made up out of the result of the Phong shading computation [147] and the sampled value. The mapping of the application onto the facade is a critical issue and choosing the right mapping depends on the shape of the underlying building. In Section 4.3, we provided mappings for the most common form factors, which we use as a foundation. There is no

```
Mtllib cube.mtl     #reference to the material library
o A_Triangle        #named object
v 0.000000 0.000000 0.000000      #vertex definition 1
v 1.000000 0.000000 0.000000      #vertex definition 2
v 1.000000 0.000000 1.000000      #vertex definition 3
v 0.000000 0.000000 1.000000      #vertex definition 4
vt 0.000000 0.000000      #2D Tex. Coord. Definition 1
vt 1.000000 0.000000      #2D Tex. Coord. Definition 2
vt 1.000000 1.000000      #2D Tex. Coord. Definition 3
vn 0.000000 1.000000 0.000000     #normal definition
usemtl Material                   #material tag
f 1/1/1 2/2/1 3/3/1               #face definition
```

Figure 5.10: The content of an *.obj* file to describe the model used in the simulator. Additional material definitions are referenced in the corresponding *.mtl* file.

canonical way to wrap a texture around a 3D surface, and reducing the dimension of the space often comes at the cost of not all edges between vertices being preserved, resulting in distortions. When mapping an application onto a facade, as depicted in Figure 5.9, an initial test image is rendered onto the facade to allow for the easy recognition of distortions. The mapping can then be adjusted manually with controls provided in the toolkit.

Model

The model describes a 3D representation of the physical properties of the building hosting the media facade, as well as its surrounding architecture. It contains objects that are present in the scenery, as well as lighting and weather conditions and any other external factors that are relevant for the scenery. This information can be included at an arbitrary level of detail. However, we do not claim to provide a full weather simulation. The toolkit offers to integrate movable light sources with different characteristics and different light intensities. They can be created and modified at run-time. With these light sources, we can simulate and test the visibility of a media facade for light of different intensity falling onto the facade from different angles and being reflected from different objects in the model. This information is also used to determine the visibility of the media facade from different places within the scene for the current conditions. Weather conditions such as fog can be directly integrated into the model when creating the model with external 3D modeling tools. The toolkit includes these information when rendering the model.

Furthermore, the model describes all areas that could possibly serve as media facade locations. In general, any surface in the 3D model could be turned into a possible location for hosting a media facade by simply naming the surface as such. When visualizing the content of an application on a media facade, the model data is used to calculate occlusion, visibility and the mapping of the application's content onto the media facade's surface.

Our toolkit supports any 3D model that is stored in the common *Wavefront*[4] **.obj** format. The basic components of this format are vertices, texture coordinates and normals. These components can be combined into faces. All faces that are not triangles are broken apart into triangles during the import of the scene. The media facade toolkit represents the entire scene as a list of model objects in the *rendering* module (see Figure 5.10) which we describe next. The model comprises a single object of the scene, containing all objects forming the scene.

Figure 5.11: An application displayed on three facades with three different resolutions. From left to right: 100%, 25% and 5% of the resolution of the original application

Rendering

The rendering module is the core module of the toolkit. It holds the main program loop, which is responsible for controlling the entire simulation process. Only one rendering module can exist within the media facade toolkit at all times. Hence, the rendering module is implemented as a *Singleton* as defined by Gamma et al. [61]. This means that only one instance of the renderer exists at any time. The rendering module contains

[4]http://goo.gl/quq8h

and manages all models, applications, facades, buffers and input senders needed for the simulation. The three main responsibilities of the renderer are:

1. When starting the simulation, the rendering module initially starts all applications that will be included in the current simulation regardless of whether they are currently mapped onto media facades. Applications that are associated with a facade are automatically rendered on the particular facade.

2. The renderer is responsible for the general rendering of the 3D scene displayed on the screen. Within this task, all buffers holding information on the scenery are rendered using appropriate shaders. The application frame buffers holding the visual content of the applications are added as textures onto the particular faces that are marked as media facades. The original resolution of the application is mapped onto the actual resolution of the particular media facade. Figure 5.11 depicts the visual appearance of an application when turning down the resolution of the facade in the toolkit. Finally, the rendering module invokes all input listeners for handling user input. With this, we can simulate media facades of arbitrary resolution.

3. While running, the rendering module continuously updates the scenery.

Hybdrid Prototyping

To allow for hybrid prototyping similar to the City Compiler [135], the toolkit allows the integration of physical miniature models into the simulation as follows (see Figure 5.12): Additional virtual cameras can be placed inside the virtual scene where they capture the virtual facade's content from their current point of view. By doing so, this virtual camera captures the facade‘s content with the correct distortion. The view rendered from the perspective of the additional cameras can be sent to external screens. We can build a physical miniature model of a building (e.g., by 3D printing) and integrate digital display elements or a pico-projector into the physical model to simulate the media facade. The position of the virtual camera relative to the virtual model matches the position of the projector or display element in the physical model relative to the physical miniature model. In this way, the media facade's content always appears correctly distorted on the physical miniature facade. To adjust the distortion of the content, the virtual camera can simply be replaced in the virtual model. By sending the output of an additional camera looking at the content of one simulated media facade to this embedded display, we create a physical model of the simulated media facade. By doing so, we allow for hybrid prototyping similar to the City Compiler [135], but with all the additional features and the flexibility provided by our toolkit.

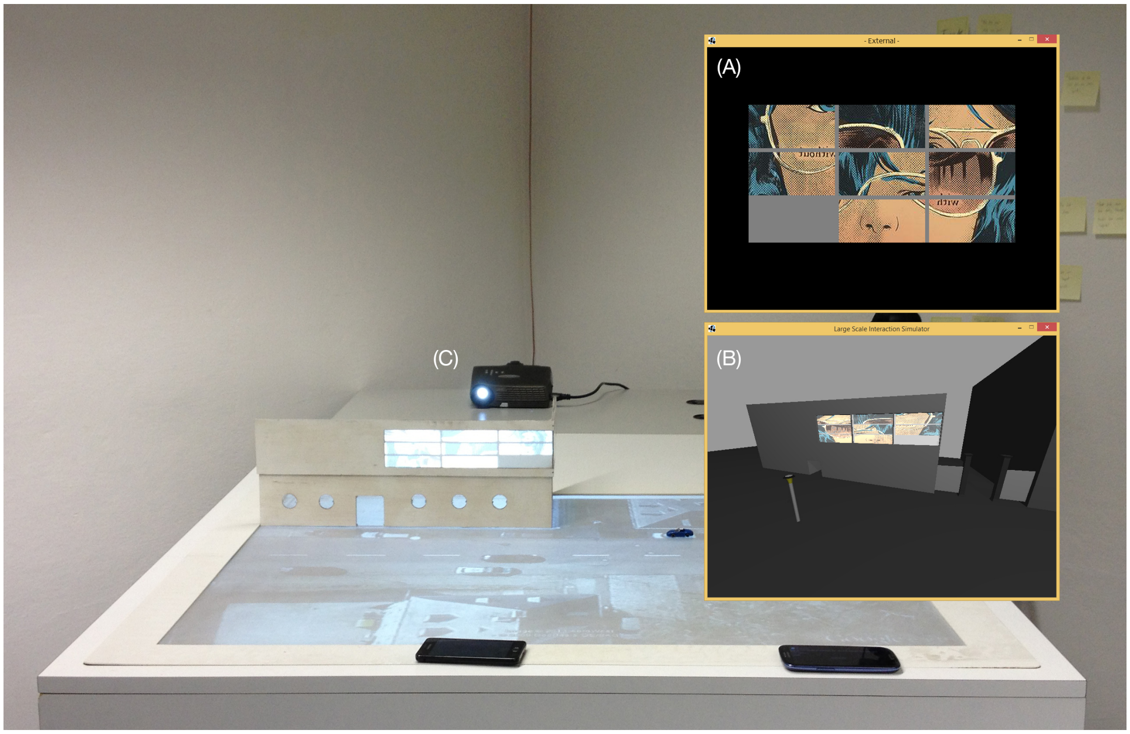


Figure 5.12: Using the toolkit's hybrid simulation feature to display the application simulated with the toolkit onto a physical miniature model of the media facade. (A) The view of the toolkit's external camera. (B) The virtual simulation of the facade. (C) Projecting the toolkit's external camera view onto a phyisical miniature model.

5.3.2 Implementation

We developed the reference implementation of the media facade toolkit using Java. Java uses a virtual machine which allows the toolkit to run on different computer architectures and operating systems without further effort. To render the 3D scenery, we utilize the "Lightweight Java Game Library"[5] (LWJGL). Through this library, we can access native OpenGL[6] functionality to render the 3D scenery directly on the GPU of the graphics card

[5]http://www.lwjgl.org
[6]http://www.opengl.org

in order to improve the performance. The applications displayed on the facade are also written in Java, implementing an interface defining methods to provide the visual content of the application to an application frame buffer and to listen for user input. The toolkit currently only supports Java applications that implement the provided interface. Applications written in other programming languages can easily be supported by providing the corresponding wrappers for the particular languages. The current implementation of the toolkit only supports a single application per facade. However, multiple applications can be supported by dividing the facade into several parts and assigning a different — or instances of the same — application to every part. In the current implementation of the toolkit, we exploit the dynamic class loading capabilities of Java to allow a developer to easily integrate new functionality and features into the toolkit on the fly. The developer only needs to provide the particular **.class** files in a specific sub-directory of the toolkit.

User interface

The media facade toolkit provides a visual representation of the rendered scene. It further provides a set of controls to create and modify the scene, as well as to add applications and input devices. The toolkit contains the following main components which are depicted in Figure 5.13: (A) The *rendering view* is the central visualization component of the toolkit. It visualizes the rendered scene and provides controls to navigate through it and to move the light source around. (B) The *control view* provides access to the settings and properties of the simulated environment. It holds the controls to create and modify the simulation environment. (C) *Control groups* organize the provided controls as collections of related controls. Controls are grouped into *Scene Graph, Mapping, Material, Facades, Applications* and *Input.* (D) The *detailed settings* provide the particular controls for the different control groups, as well as additional information about the current settings.

The user interface of the toolkit is intentionally separated into a rendering view and a control view. The rendering can be displayed on a separate screen. By utilizing a screen larger than a common desktop screen, such as a projection wall, we can create a more holistic impression of the simulated environment and provide means for a more immersive and realistic interaction.

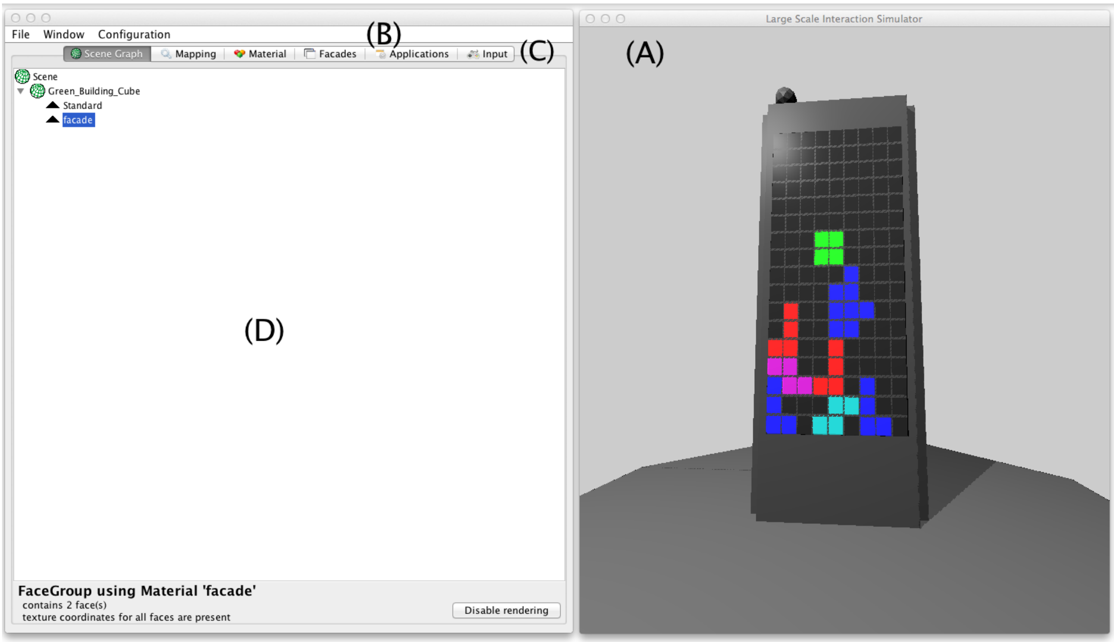


Figure 5.13: The interface of the media facade toolkit: (A) the rendering view, (B) the control view, (C) control groups and (D) the detailed settings.

5.3.3 User Feedback

In order to gather feedback on the practicability of the toolkit, we distributed it to five media facade application developers to utilize it for prototyping and simulating interactive applications for media facades (see Figure 5.14). Developers *A* and *B* are media informatics students, developers *C* and *D* are media facade application developers and developer *E* is a media facade professional involved in building large-scale media facades. In this qualitative evaluation, our interest was in getting feedback on how the media facade toolkit performs in real settings and how it can enhance the process of designing transferable, interactive applications for media facades and overcome limitations of current approaches. We repeatedly conducted interviews with the users during the particular stages of the design process. The developers were not involved in the development process of the media facade toolkit. Developers C and D implemented interactive applications that were deployed on an actual media facade. Developers A, B and E developed standalone applications, adapted them to run within the toolkit and visualized them on several virtual media facades with different form factors. All of them already had experience in developing interactive applications for media facades and they are thus familiar

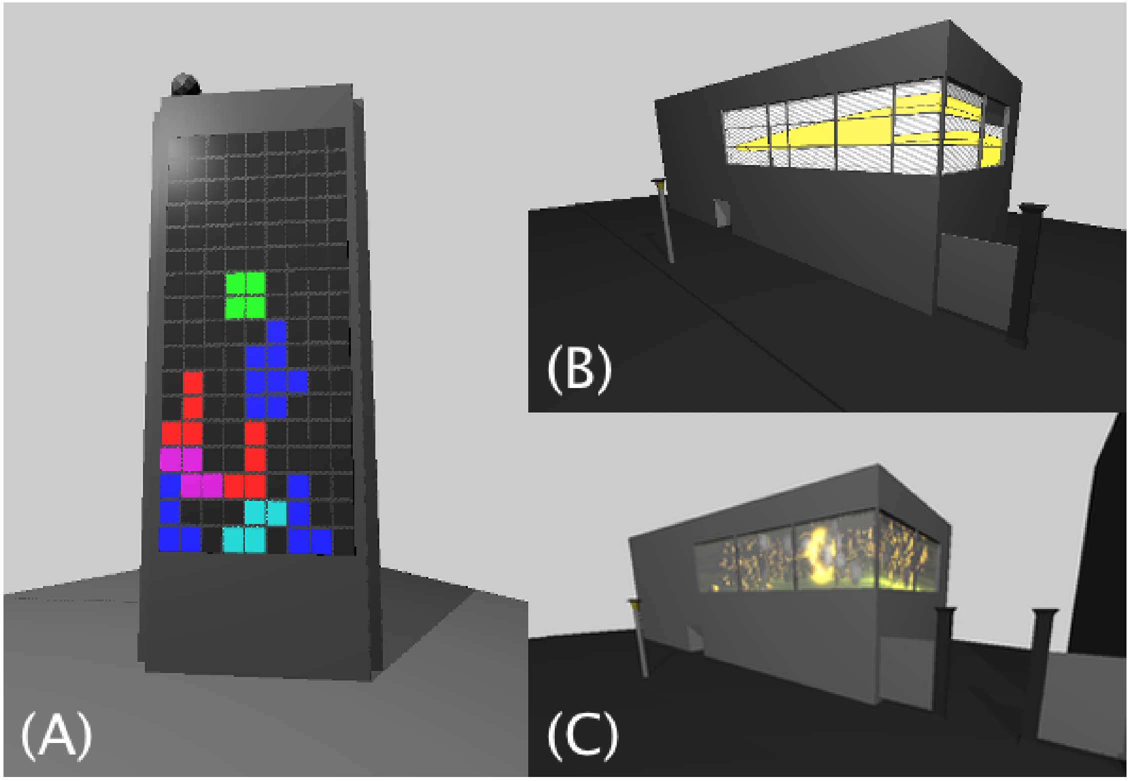


Figure 5.14: Applications developed using our toolkit. (a) Tetris: This game was developed as a standalone application and simulated on differently shaped facades. (b) Move-the-tile: A multi-user puzzle. The application was developed using the facade toolkit throughout the design process. (c) A video player to display videos and animation on various media facades.

with current design approaches and their pitfalls. They all stated that using our toolkit indeed simplified the development process tremendously. In particular, to demonstrate the benefits of the media facade toolkit, we want to outline one example of how the toolkit was used: Within the scope of conducting a user study for a project investigating multi-user interaction with a shared media facade, developer D utilized the media facade toolkit to develop a multi-user puzzle game intended to run on differently shaped displays in different settings and locations. The game was displayed on a miniature model of a media facade using the toolkit's hybrid prototyping feature, as well as two different media facades of different size, shape and technology. In current practice, such a scenario would require extensive modifications of the application (e.g., re-writing large parts of the code) to run in all three settings. Using our toolkit, developer D implemented the application as a standard Java application, without addressing any particular facade. Running the developed application with the media facade toolkit involved the following five steps:

1. Loading the particular 3D model of the facade into the toolkit. In general, if it is not available, it can be easily created with any common modeling software.

2. Selecting the surface representing the media facade in the 3D model as well as an appropriate mapping from the set of available mappings.

3. Assigning an application to the facade and

4. selecting an input mapping.

5. The application can be started. It is automatically mapped onto the media facade in the 3D model and the graphical output is additionally sent to the connected display or facade.

Once it is running in the toolkit, transferring the application to a different facade only requires repeating steps (1) and (2). After finishing the deployment of his application in the three different settings, we asked developer D about his experiences with our toolkit in a semi-structured interview. He stated that he liked being able to transfer an application to another facade without modifying the application itself. Within this scope, he also liked the possibility of getting immediate visual feedback on how the application will be displayed on different facades in the toolkit's simulation. He mentioned the transferability of applications as the main benefit for his project. Furthermore, he liked the option to reuse the interaction client he developed in his project with arbitrary applications running in the toolkit.

From developer E, we received valuable feedback from a professional perspective. He mentioned the re-usability of existing applications as a positive aspect while pointing out the possibility of prototyping and simulating complex installations involving various form factors and technologies as the main benefit, since this was the most time-consuming and costly part of the development process. He further stated that in current practice, complex settings need to be rebuilt with real hardware as abstracted models for prototyping and testing and only a fraction of the work spent on this can be incorporated in the final result. In addition, he complained that this approach of current prototyping does not provide a good impression of the overall installations, since it only models certain features. He also liked the possibility of modeling the surroundings of the building to simulate different scenarios. Asked about the simulation component of the toolkit, he liked being able to have a holistic visual representation of the whole scenery — including the media facade as well as it's surroundings — which allows for dynamically changing the current point of view and which provides immediate visual feedback on changes.

In general, all users were able to develop their applications as standalone applications without immediately tailoring them to a specific facade. None of the developers needed more than two hours to adapt the completed application to run on a particular facade. Once it was running on a facade within the toolkit, it took them on average less than five minutes to load the 3D model of a different facade and to transfer the application from the current facade to another one with a different form factor. A request made by two users was to provide general-purpose client applications to allow for controlling a pointer on the facade in order to take the burden of developing them away from the designers. This is an issue we want to address in future work.

5.4 Discussion

We introduced a rapid prototyping toolkit which allows for the development and simulation of interactive applications on arbitrary media facades. Due to its modular design, it provides a high degree of flexibility. The toolkit allows for the smooth combination of various applications, interaction techniques, media facades and building models where every module can be exchanged with ease. This offers the potential to quickly adapt to new settings. Hence, the media facade toolkit shows a high level of potential to become a universal design tool for the designers and developers of digital content for media facades and other urban screens. In contrast to existing toolkits, a designer can easily try out given content on different media facades, in various locations, without much effort. An application can be developed without tailoring it to a specific facade or to a specific technical setup. It can be developed as a standalone application which can be mapped onto arbitrary media facades with the media facade toolkit. Furthermore, the toolkit can also be used as a standalone visualization tool. General visual content can be simulated with the toolkit as well, including pre-produced videos and animations.

Despite the potential to ease the development of interactive applications for media facades and to develop applications that can be immediately ported to different facade types, the media facade toolkit also comes with certain limitations. Due to the technical and physical properties of media facades and their highly dynamic, public deployment context, it is hardly possible to cover all settings to a full extent. Our work instead aims to provide a flexible prototyping and simulation framework into which arbitrary building models, media facades, applications and interaction techniques can be integrated in order to easily adapt existing setups and to ensure the portability of applications from one media facade to another. For the sake of simplicity, we currently only address pixel-based media facades since they represent the vast majority of media facades. Mechanical media facades — media facades that are created by physically movable mechanical elements of a building — and voxel facades are currently not covered. Their implementation requires dynamic geometry, a feature that is currently not supported by the toolkit. Furthermore, automatically choosing an appropriate texture mapping is not supported by the toolkit. Since existing media facades have various different form factors, this remains an open problem. As previously mentioned, the media facade toolkit currently limits the number of applications per facade to one. This limitation can be easily circumvented by dividing the facade into multiple facades which can display one application each.

6

Conclusion

The goal of this thesis was to exploit the capabilities of media facades in urban environments in order to take the first step towards turning media facades into fully fledged interactive surfaces. With this, we offer new perspectives to the field of urban computing by providing means to utilize media facades as visual interfaces between inhabitants and elements of the urban environment (e.g., buildings, infrastructure, etc.). In this section, we summarize the work presented in this thesis, and we outline the achieved contributions. This is followed by identifying directions for future work. We conclude this thesis with some closing remarks on the potential of urban media facades and media architecture.

6.1 Summary

The work presented in this thesis follows three directions. After reviewing work from the areas of shared encounters in public spaces, interaction with public screens of various sizes as well as designing and prototyping interaction for public spaces in Chapter 2, we investigated interaction with media facades more holistically on a theoretical level in Chapter 3. From the results of this investigation, we derived taxonomies to allow a categorization and comparison of interactive media facade installations based on the resolution of the facade and the applied input techniques, as well as the technical realization of the installation as a whole, addressing characteristic properties. In Chapter 4, we developed three techniques to interact with media facades without the need for instrumenting the potential interactive space in front of the facade. We provided techniques to interact with media facades of

1. various resolutions, specifically addressing low-resolution media facades,
2. media facades with arbitrary form factors and
3. interaction by pointing at the facade with a mobile device.

All three techniques utilize the personal mobile devices of the users. Finally, in Chapter 5 we developed a generalized media facade toolkit for rapid prototyping and simulating interactive installations for media facades, which is capable of simulating large-scale media facade installations. The simulation is independent of the size, form factor, technology and hardware of the simulated facade. The toolkit represents a flexible simulation, test and development tool for designers and researchers alike, addressing the diversity of media facade settings and incorporating the media facade's architectural surroundings and the deployment space.

6.1.1 Contributions

After summarizing the work conducted in this thesis, we now outline the achieved contributions. With the rapidly growing number of media facades in urban environments, their role as digital screens will gain in importance. In the light of the unique characteristics of media facades and their deployment spaces, we addressed the problem of how to allow for interactivity with media facades beyond playful arts-related installations. This thesis contributes to exploiting the potential of media facades to turn them into the future interactive computing surfaces of urban spaces in three ways:

- We provided a theoretical analysis of interaction with media facades,
- novel interaction techniques to interact with media facades at a distance, as well as
- prototyping and simulation tools to allow for the generalized development of interactive installations for media facades.

Next, we outline the achieved contributions of this thesis in more detail, with respect to their role as theoretical, technical and design contributions:

1. **Theoretical contribution:** The lack of a theoretical understanding of which parameters define and shape interaction with media facades raises various challenges that designers an developers have to face. Furthermore, there is neither a common ground on the suitability of different interaction techniques in particular settings,

nor is there any categorization that puts the components of media facade installations into context. We contributed to this by analyzing the design space for interactive installations for media facades on a theoretical level. We investigated the extent to which an interactive installation allows passersby to participate, the resolution of the media facade itself with respect to the resolution of user input, as well as the technical characteristics of the overall setting. We provided taxonomies for each of these three aspects, which put media facade installations in context and allow for comparing them with each other. Furthermore, we introduced measures accompanying the taxonomies: First, a resolution factor that relates input and facade resolution to indicate the suitability of an interaction technique to allow for a particular granularity of interaction for a media facade. Second, a set of observations and guidelines to derive properties of the interaction from the technical characteristics of an interactive media facade installation, as well as to determine if an interaction technique is applicable in a given setting (see Chapter 3).

2. **Technical contribution:** Due to their size and the therefore required minimal viewing distance, media facades usually demand interaction at a distance. In order to interact with media facades as interactive surfaces, we need interaction techniques allowing for fine-grained interaction with the media facade's content with respect to its resolution and shape. We contributed to this by providing three novel interaction techniques allowing for fine-grained interaction with a media facade. All three techniques allow multiple users to interact simultaneously. They utilize regular mobile devices and do not require an instrumentation of the potential interactive space around the media facade with additional technology. This makes them applicable in a wide range of settings. They further address the form factor and the resolution of the facade in particular as characteristic properties. We provided a technique to interact with media facades of arbitrary resolution by allowing for interaction through live video on mobile devices (see Section 4.1). This approach contributes especially to interacting with low-resolution facades, since visual feedback can be augmented in the live video on the mobile device without wasting screen real estate on the facade. Additionally, we provided a standalone client applying the provided technique, which we contributed to the TUIO community. With the GPS Lens, we provided a sensor-based interaction technique to control a pointer on a media facade by pointing at the facade with a mobile device (see Section 4.2). Additionally, we provided a technique that particularly addresses the form factor of a media

facade. With facade maps, we provided a technique to interact with media facades of arbitrary form factors using cartographic map projections on a mobile device, where users interact with a cartographic map of the media facade's surface. Within this scope, we further derived common facade shapes from reviewing existing media facades and we provided map projections suitable for the particular shape (see Section 4.3).

3. **Design contribution:** The technical aspects of media facades, the context in which they are deployed and the exposure of their content to a large audience raise the need for a tailored design in terms of design and prototyping tools. We contributed to the design and development of interactive media facade installations by providing a generalized media facade toolkit for rapid prototyping and simulating interactive media facade installations. In particular, the toolkit can simulate large-scale media facades, independent of the media facade's size, form factor, technology and underlying hardware. It further makes it possible to run arbitrary applications on arbitrary media facades, reducing the strong bond between application and media facade that exists with current development toolkits. With this, we allow for the transferability of applications from one media facade to another, which has so far been one of the major disadvantages of current development tools. Applications can be mapped onto media facades as texture images by integrating cartographic map projections into the toolkit to determine texture coordinates for arbitrarily shaped surfaces. With the media facade toolkit, we further allow simulation of different hardware, media facade resolutions and different color models. It further addresses the diversity of media facade settings and incorporates the media facade's architectural surroundings into the simulation (see Chapter 5).

6.2 Future Work

With the work presented in this thesis, we provided new perspectives for interacting with media facades and designing interactive media facade installations alike. With our work, we laid the groundwork for utilizing media facades as interactive surfaces in urban environments. We further provided starting points for future research based on the contributions of this thesis. For future research, we identified the following main directions:

Multi-User Interaction

When multiple users interact with a shared screen, different issues can arise. First, when allowing multiple users to access a shared media facade simultaneously, territorial issues usually arise. Users compete for access to a shared screen, which can lead to frustration and diminish the user experience (see Section 4.1). The usually enormous size and great visibility of media facades in general amplify this issue since they allow for a relatively large number of possible users, sharing limited screen real estate. Hence, we need to further investigate multi-user interaction concepts for media facades, identifying possible display sharing strategies to allow for a balanced interaction. In addition to sharing the screen, as for example by taking turns or assigning exclusive regions to the particular users, simultaneous interaction with shared content is an important aspect that needs to be investigated. While taking turns and sub-dividing the screen decrease the available screen real estate or restrict access to the media facade, by mediating simultaneous access on a shared media facade, all users can interact, exploiting the media facade's characteristics.

A further aspect that needs further investigation in the future is the awareness of other users. As we found out, being unaware of other users can lead to competitive scenarios that lead to frustration and it can diminish the user experience when interacting with a media facade. To avoid this, we need to investigate ways to make users aware of each other. Within this scope, using a media facade as a means to engage and mediate interaction between users also is a promising path.

Standardization

With this thesis, we provided a first approach to developing generalized interaction techniques to control pointers and manipulate content of arbitrary media facades without requiring the instrumentation of the potential interaction space around the facade. To take this further, we need to develop standardized protocols for user input, such as the TUIO protocol for multi-touch, as well as for exchanging content with a media facade. As pointed out in this thesis, current development and simulation tools for media facades are strongly connected with the one particular facade they were created for. To ease the development process, there is a need for creating generalized development tools that are detached from any particular facade. So far, there is no common ground for approaching the development of interaction with media facades in a generalized way. We need to further investigate the process of designing and developing interactive media facade installations in order to derive suitable design processes and evaluation methods tailored to the characteristics of media facades. Additionally, we need to extend and refine the taxonomies presented in this thesis. We have to deepen our understanding of what the driving parameters for interactive media facade installations are and improve the taxonomies to include them and to provide an even more fine-grained categorization.

Prototyping

We provided a first generalized prototyping and simulation toolkit, which allows for designing interactive media facade installations without binding the application to a particular facade. However, in future work, we plan to extend the media facade toolkit in several ways. Firstly, we need to introduce additional features such as multiple — also moving — light sources (e.g. cars that drive by, light emitted from neighboring displays or street lights, etc.) in the environment around the media facade, simulating passersby dynamically in the scenery to increase the level of detail in which a media facade and its surrounding environment can be mapped. Within this scope, we also want to integrate live data from the real-world counterpart of a simulated media facade environment, to simulate realistic settings. Since the current version of the toolkit only supports uniformly shaped and distributed pixels, we further want to allow for different pixel shapes. Additionally, establishing compatibility with other simulation tools and toolkits is an important aspect for supporting existing media facade installations in the toolkit. Along with this, the integration of the media facade toolkit with existing media facades and their interfaces with respect to pushing content onto a media facade using the toolkit needs to be investigated further.

Users

This thesis investigated interaction with media facades from the perspective of how to create interactive media facade installations in general beyond an artistic or playful context, how to design and how to describe such interactive media facade installations. With the contributions of our work, we lay the foundations for new forms of interaction in urban spaces. Within this scope, we need to illuminate the interaction with media facades from a more user-driven perspective. As previously pointed out, the highly public context in which the interaction occurs, influences both the way people interact with a shared media facade and with each other, as well their experiences while interacting. Within this scope, in order to investigate the users' actions and emotions more holistically, we need to develop suitable evaluation methods and adapt existing ones to the context of media facades. To substantiate their validity, we need to apply these concepts involving different scenarios, different technologies and interaction methods. This may lead to more profound knowledge regarding which form of interaction is appropriate for achieving a particular experience. With that, design processes in such a context could, in future setups, directly be driven by aiming at certain experience-related goals.

6.3 Closing Remarks

Over the course of this thesis, we holistically investigated interaction with media facades. With the theoretical analysis, the novel interaction techniques and the provided prototyping toolkits, we set the starting point for the generalized and structured development of interactive media facade installations. To further substantiate our work with the help of the community, we publish relevant parts of our findings as open source. With the work described in this thesis and the lessons learned through the design and deployment of our solutions, we want to inspire others to take up our approach and enhance the development of interactivity for media facades to advance our goal of turning media facades into the future interactive computing surfaces of urban environments.

List of Figures

List of Tables

Bibliography

[1] M. Ardito. Studies of the Influence of Display Size and Picture Brightness on the Preferred Viewing Distance for HDTV Programs. *Motion Imaging Journal, SMPTE*, 103(8):517–522, 1994.

[2] A. Azad, J. Ruiz, D. Vogel, M. Hancock, and E. Lank. Territoriality and behaviour on and around large vertical publicly-shared displays. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, pages 468–477, New York, NY, USA, 2012. ACM.

[3] R. Balakrishnan and P. Baudisch. Special issue on ubiquitous multi-display environments. *HumanComputer Interaction*, 24(1-2):1–8, 2009.

[4] R. Ballagas, J. Borchers, M. Rohs, and J. Sheridan. The smart phone: a ubiquitous input device. *Pervasive Computing, IEEE*, 5(1):70–77, 2006.

[5] R. Ballagas, M. Rohs, and J. G. Sheridan. Sweep and point and shoot: phonecam-based interactions for large public displays. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1200–1203, New York, NY, USA, 2005. ACM.

[6] T. Ballendat, N. Marquardt, and S. Greenberg. Proxemic interaction: designing for a proximity and orientation-aware environment. In *ACM International Conference on Interactive Tabletops and Surfaces*, ITS '10, pages 121–130, New York, NY, USA, 2010. ACM.

[7] A. Banerjee, J. Burstyn, A. Girouard, and R. Vertegaal. MultiPoint: comparing laser and manual pointing as remote input in large display interactions. *International Journal of Human-Computer Studies*, 70(10):690 – 702, 2012. Special issue on Developing, Evaluating and Deploying Multi-touch Systems.

[8] J. A. Bargas-Avila and K. Hornbaek. Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2689–2698, New York, NY, USA, 2011. ACM.

[9] J. Bartlett. Rock 'n' scroll is here to stay [user interface]. *Computer Graphics and Applications, IEEE*, 20(3):40–45, 2000.

[10] P. Baudisch, M. Sinclair, and A. Wilson. Soap: a pointing device that works in mid-air. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology*, UIST '06, pages 43–46, New York, NY, USA, 2006. ACM.

[11] T. Bauer. *Designing Multi-User Interaction with Media Façades*. Ludwig-Maximilians-Universität München, Department of Computer Science, 2013.

[12] D. Baur, S. Boring, and S. Feiner. Virtual projection: exploring optical projection as a metaphor for multi-device interaction. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems*, CHI '12, pages 1693–1702, New York, NY, USA, 2012. ACM.

[13] J. Baus, A. Krüger, and W. Wahlster. A resource-adaptive mobile navigation system. In *Proceedings of the 7th International Conference on Intelligent User Interfaces*, IUI '02, pages 15–22, New York, NY, USA, 2002. ACM.

[14] W. Beer. GeoPointer: approaching tangible augmentation of the real world. In *Proceedings of the 8th International Conference on Advances in Mobile Computing and Multimedia*, MoMM '10, pages 221–225, New York, NY, USA, 2010. ACM.

[15] G. Beyer, F. Köttner, M. Schiewe, I. Haulsen, and A. Butz. Squaring the circle: how framing influences user behavior around a seamless cylindrical display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1729–1738, New York, NY, USA, 2013. ACM.

[16] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose. Toolglass and magic lenses: the see-through interface. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, pages 73–80, New York, NY, USA, 1993. ACM.

[17] E. A. Bier, M. C. Stone, K. Pier, K. Fishkin, T. Baudel, M. Conway, W. Buxton, and T. DeRose. Toolglass and magic lenses: the see-through interface. In *Conference Companion on Human Factors in Computing Systems*, CHI '94, pages 445–446, New York, NY, USA, 1994. ACM.

[18] M. Böhmer, S. Gehring, M. Löchtefeld, M. Ostkamp, and G. Bauer. The mighty un-touchables: creating playful engagement on media façades. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '11, pages 605–610, New York, NY, USA, 2011. ACM.

[19] R. A. Bolt. *"Put-that-there": Voice and Gesture at the Graphics Interface*, volume 14:3. ACM, 1980.

[20] S. Boring. *Interacting "Through the Display"*. Ludwig-Maximilians-Universität München, Juli 2010.

[21] S. Boring, M. Altendorfer, G. Broll, O. Hilliges, and A. Butz. Shoot & copy: phonecam-based information transfer from public displays onto mobile phones. In *Proceedings of the 4th International Conference on Mobile Technology, Applications, and Systems and the 1st International Symposium on Computer Human Interaction in Mobile Technology*, Mobility '07, pages 24–31, New York, NY, USA, 2007. ACM.

[22] S. Boring and D. Baur. Making public displays interactive everywhere. *Computer Graphics and Applications, IEEE*, 33(2):28–36, 2013.

[23] S. Boring, D. Baur, A. Butz, S. Gustafson, and P. Baudisch. Touch Projector: mobile interaction through video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2287–2296, New York, NY, USA, 2010. ACM.

[24] S. Boring, S. Gehring, A. Wiethoff, A. M. Blöckner, J. Schöning, and A. Butz. Multi-user interaction on media facades through live video on mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2721–2724, New York, NY, USA, 2011. ACM.

[25] S. Boring, M. Jurmu, and A. Butz. Scroll, tilt or move it: using mobile phones to continuously control pointers on large public displays. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, OZCHI '09, pages 161–168, New York, NY, USA, 2009. ACM.

[26] F. Borries, S. Walz, and M. Böttger. *Space Time Play: Computer Games, Architecture and Urbanism: The Next Level*. Birkhüser, Basel, 2007.

[27] D. Bouchard. *Embodied Emergence: Distributed Computing Manipulatives*. PhD thesis, Massachusetts Institute of Technology. Dept. of Architecture. Program in Media Arts and Sciences., 2007.

[28] H. Brignull and Y. Rogers. Enticing people to interact with large public displays in public spaces. In M. Rauterberg, M. Menozzi, and J. Wesson, editors, *Proceedings of the IFIP International Conference on Human-Computer Interaction*, Interact'03, pages 17–24. IOS Press, 2003.

[29] G. Broll, R. Graebsch, M. Scherr, S. Boring, P. Holleis, and M. Wagner. Touch to play – exploring touch-based mobile interaction with public displays. In *Near Field Communication (NFC), 2011 3rd International Workshop on*, pages 15–20, 2011.

[30] M. Brynskov, P. Dalsgaard, T. Ebsen, J. Fritsch, K. Halskov, and R. Nielsen. Staging urban interactions with media façades. In T. Gross, J. Gulliksen, P. Kotz, L. Oestreicher, P. Palanque, R. Prates, and M. Winckler, editors, *Human-Computer Interaction INTERACT 2009*, volume 5726 of *Lecture Notes in Computer Science*, pages 154–167. Springer Berlin Heidelberg, 2009.

[31] M. Buchenau and J. F. Suri. Experience prototyping. In *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, DIS '00, pages 424–433, New York, NY, USA, 2000. ACM.

[32] L. Bullivant. *Responsive Environments: Architecture, Art and Design.* V&A, London, 2006.

[33] M. Burmester, M. Mast, K. Jäger, and H. Homans. Valence method for formative evaluation of user experience. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, DIS '10, pages 364–367, New York, NY, USA, 2010. ACM.

[34] W. Buxton. *Sketching User Experiences: Getting the Design Right and the Right Design.* Morgan Kaufmann, 2007.

[35] W. Buxton. *User Experiences: Getting the Design Right and the Right Design.* Morgan Kaufmann, 2007.

[36] J. Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PAMI-8(6):679–698, 1986.

[37] A. Caragliu, C. Del Bo, and P. Nijkamp. Smart Cities in Europe. Serie Research Memoranda 0048, VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics, 2009.

[38] J. B. Carroll. *Human Cognitive Abilities: A Survey of Factor-Analytic Studies.* Cambridge University Press, 1993.

[39] E. Catmull. *A Subdivision Algorithm for Computer Display of Curved Surfaces.* University of Utah, 1974.

[40] T.-H. Chang and Y. Li. Deep shot: a framework for migrating tasks across devices using mobile phone cameras. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 2163–2172, New York, NY, USA, 2011. ACM.

[41] K. Cheverst, A. Dix, D. Fitton, C. Kray, M. Rouncefield, C. Sas, G. Saslis-Lagoudakis, and J. G. Sheridan. Exploring Bluetooth based mobile phone interaction with the Hermes Photo Display. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services*, MobileHCI '05, pages 47–54, New York, NY, USA, 2005. ACM.

[42] S.-J. Cho, C. Choi, Y. Sung, K. Lee, Y.-B. Kim, and R. Murray-Smith. Dynamics of tilt-based browsing on mobile devices. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '07, pages 1947–1952, New York, NY, USA, 2007. ACM.

[43] R. Dachselt and R. Buchholz. Natural throw and tilt interaction between mobile phones and distant displays. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, pages 3253–3258, New York, NY, USA, 2009. ACM.

[44] P. Dalsgaard, C. Dindler, and K. Halskov. Understanding the dynamics of engaging interaction in public spaces. In P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, and M. Winckler, editors, *Human-Computer Interaction INTERACT 2011*, volume 6947 of *Lecture Notes in Computer Science*, pages 212–229. Springer Berlin Heidelberg, 2011.

[45] P. Dalsgaard and K. Halskov. Designing urban media façades: cases and challenges. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2277–2286, New York, NY, USA, 2010. ACM.

[46] P. Dalsgaard and K. Halskov. Reflective design documentation. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, pages 428–437, New York, NY, USA, 2012. ACM.

[47] P. Dalsgaard, K. Halskov, and R. Nielsen. Towards a design space explorer for media facades. In *Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat*, OZCHI '08, pages 219–226, New York, NY, USA, 2008. ACM.

[48] J. Davis and X. Chen. Lumipoint: multi-user laser-based interaction on large tiled displays. *Displays*, 23(5):205 – 211, 2002.

[49] P. Dietz and D. Leigh. DiamondTouch: a multi-user touch technology. In *Proceedings of the 14th Annual ACM Aymposium on User Interface Software and Technology*, UIST '01, pages 219–226, New York, NY, USA, 2001. ACM.

[50] N. V. Diniz, C. A. Duarte, and N. M. Guimarães. Mapping interaction onto media façades. In *Proceedings of the 2012 International Symposium on Pervasive Displays*, PerDis '12, pages 14:1–14:6, New York, NY, USA, 2012. ACM.

[51] A. Dix, J. Finlay, G. D. Abowd, and R. Beale. *Human Computer Interaction*. Prentice Hall, 2003.

[52] P. Dourish. Re-space-ing place: "place" and "space" ten years on. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work*, CSCW '06, pages 299–308, New York, NY, USA, 2006. ACM.

[53] E. Eriksson, T. R. Hansen, and A. Lykke-Olesen. Reclaiming public space: designing for public interaction with private devices. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, TEI '07, pages 31–38, New York, NY, USA, 2007. ACM.

[54] P. T. Fischer and E. Hornecker. Urban HCI: spatial aspects in the design of shared encounters for media facades. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 307–316, New York, NY, USA, 2012. ACM.

[55] P. T. Fischer, C. Zöllner, T. Hoffmann, and S. Piatza. VR/Urban: SMSlingshot. In *Proceedings of the 4th International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '10, pages 381–382, New York, NY, USA, 2010. ACM.

[56] P. T. Fischer, C. Zöllner, and E. Hornecker. VR/Urban: Spread.gun - design process and challenges in developing a shared encounter for media façades. In *Proceedings of the 24th BCS Interaction Specialist Group Conference*, BCS '10, pages 289–298, Swinton, UK, UK, 2010. British Computer Society.

[57] J. D. Foley. *Computer Graphics: Principles and Practice*. Addison-Wesley Professional, 1996.

[58] J. Forlizzi and K. Battarbee. Understanding experience in interactive systems. In *Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, DIS '04, pages 261–268, New York, NY, USA, 2004. ACM.

[59] D. Frenchman and F. Rojas. Zaragoza's Digital Mile: Place-Making in a New Public Realm. *Places*, 18(2), 16, 2006.

[60] J. Fritsch and P. Dalsgaard. Media Façades beyond interaction. In *Proceedings of the 20th Australasian Computer-Human Interaction Conference*, OZCHI'08, pages 8 – 12. ACM, 2008.

[61] E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design Patterns*. Addison-Wesley, 1995.

[62] J. Gehl. *Life Between Buildings: Using Public Space*. Island Press, 6th edition edition, 2011.

[63] S. Gehring. Media facades: turning buildings into large-scale interactive surfaces. In *Proceedings of the Workshop on "Displays Take New Shape: An Agenda for Future Interactive Surfaces". ACM International Conference on Human Factors in Computing Systems (CHI-13), Changing Perspectives, April 27 - May 2, Paris, France*. ACM, 2013.

[64] S. Gehring, F. Daiber, and C. Lander. Towards universal, direct remote interaction with distant public displays. In *Proceedings of the 3rd Workshop on Infrastructure and Design Challenges of Coupled Display Visual Interfaces (Capri, Italy), PPD*, volume 12, page 302, 2012.

[65] S. Gehring, E. Hartz, M. Löchtefeld, and A. Krüger. The media façade toolkit: prototyping and simulating interaction with media façades. In *Proceedings of the*

2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp '13, pages 763–772, New York, NY, USA, 2013. ACM.

[66] S. Gehring and A. Krüger. Façade map: continuous interaction with media façades using cartographic map projections. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 471–480, New York, NY, USA, 2012. ACM.

[67] S. Gehring and A. Krüger. Using media façades to engage social interaction. In *Adjunct Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 794–795, New York, NY, USA, 2012. ACM.

[68] S. Gehring and C. Lander. GPS Lens: GPS based controlling of pointers on large-scale urban displays using mobile devices. In *Proceedings of the 2nd ACM International Symposium on Pervasive Displays*, PerDis '13, pages 115–120, New York, NY, USA, 2013. ACM.

[69] S. Gehring and A. Wiethoff. Digital light installations - connecting people through interactive buildings. In *Workshop on Interactive City Lighting. ACM International Conference on Human Factors in Computing Systems (CHI-13), Changing Perspectives, April 27 - May 2, Paris, France*. ACM, 4 2013.

[70] D. H. Gelernter. *Mirror Worlds, or, The Day Software Puts the Universe in a Shoebox: How It Will Happen and What It Will Mean.* Oxford University Press New York, 1991.

[71] A. F. gen. Schieck, V. Kostakos, and A. Penn. Exploring digital encounters in the public arena. In K. S. Willis, G. Roussos, K. Chorianopoulos, and M. Struppek, editors, *Shared Encounters*, 1431-1496, pages 179–195, London, UK, 2020. Springer.

[72] A. F. gen. Schiek, C. Briones, and C. Mottram. The urban screen as a socialising platform: exploring the role of place within the urban space. In F. Eckardt, J. Geelhaar, L. Colini, K. S. Willis, K. Chorianopoulos, and R. Hennig, editors, *MEDIACITY: Situations, Practices and Encounters*, MEDIACITY'08, pages 285–305, Berlin, Germany, 2008. Frank & Timme GmbH.

[73] A. Geven, J. Schrammel, and M. Tscheligi. Narrations and storytelling as methodological key elements for studying user experience. In *Proceedings of NordiCHI'06*. ACM, 2006.

[74] R. Giffinger, C. Fertner, H. Kramar, R. Kalasek, N. Pichler-Milanovic, and E. Meijers, editors. *Smart Cities - Ranking of European Medium-Sized Cities.* Vienna Centre of Regional Science, http://www.smart-cities.eu/, 2007.

[75] E. Goffman. *Behavior in Public Places - Notes on the Social Organization on Gatherings*. The Free Press, 1963.

[76] E. W. Grafarend and F. W. Krumm, editors. *Map Projections: Cartographic Information Systems*. Springer, 2006.

[77] S. Greenberg and C. Fitchett. Phidgets: easy development of physical interfaces through physical widgets. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology*, UIST '01, pages 209–218, New York, NY, USA, 2001. ACM.

[78] A. Greenfield and M. Shepard. *Urban Computing and its Discontents.* The Architectural League of New York, 2007.

[79] D. Greenhood. *Mapping.* University of Chicago Press, 1964.

[80] C. Guitiérrez, M. Canet, J. Puig, and J. Lloret. Lummoblocks. `http://www.lummo.eu/lummotetris.html`. Accessed: 2013-06-30.

[81] H. M. Haeusler. *Media Facades - History, Technology, Content.* avedition, 2009.

[82] H. M. Haeusler, M. Tomitsch, and G. Tscherteu. *New Media Facades - A Global Survey.* avedition, 2013.

[83] K. Halskov and T. Ebsen. A framework for designing complex media facades. *Design Studies*, 2013.

[84] T. Hansen, E. Eriksson, and A. Lykke-Olesen. Mixed interaction space - expanding the interaction space with mobile devices. In T. McEwan, J. Gulliksen, and D. Benyon, editors, *People and Computers XIX - The Bigger Picture*, pages 365–380. Springer London, 2006.

[85] T. R. Hansen, E. Eriksson, and A. Lykke-Olesen. Mixed interaction space: designing for camera based interaction with mobile devices. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1933–1936, New York, NY, USA, 2005. ACM.

[86] R. Hardy and E. Rukzio. Touch & interact: touch-based interaction of mobile phones with displays. In *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '08, pages 245–254, New York, NY, USA, 2008. ACM.

[87] R. Hardy, E. Rukzio, P. Holleis, and M. Wagner. Mobile interaction with static and dynamic NFC-based displays. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '10, pages 123–132, New York, NY, USA, 2010. ACM.

[88] R. Hardy, E. Rukzio, M. Wagner, and M. Paolucci. Exploring Expressive NFC-Based Mobile Phone Interaction with Large Dynamic Displays. In *Near Field Communication, 2009. NFC '09. First International Workshop on*, pages 36–41, 2009.

[89] S. Harrison and P. Dourish. Re-place-ing space: the roles of place and space in collaborative systems. In *Proceedings of the 1996 ACM Conference on Computer*

Supported Cooperative Work, CSCW '96, pages 67–76, New York, NY, USA, 1996. ACM.

[90] E. Hartz. *LASIS - A Simulator for Media Façade Interaction.* Saarland University, Saarbrücken, Germany, Department of Computer Science, 2013.

[91] M. Hassenzahl. Experience Design: Technology for All the Right Reasons. *Synthesis Lectures on Human-Centered Informatics*, 3, 2010.

[92] M. Hassenzahl, S. Diefenbach, and A. Göritz. Needs, affect, and interactive products - facets of user experience. *Interacting with Computers*, 22, 2010.

[93] P. Heckbert. A survey of texture mapping. *IEEE Computer Graphics and Applications*, 6(11):56–67, 1986.

[94] C. E. Herr. Non-trivial media façades. In *Beyond Codes and Pixels: Proceedings of the 17th International Conference on Computer-Aided Architectural Design Research in Asia*, pages 99–108. Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Hong Kong, 2012.

[95] E. Hornecker, P. Marshall, and Y. Rogers. From entry to access: how shareability comes about. In *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces*, DPPI '07, pages 328–342, New York, NY, USA, 2007. ACM.

[96] P. Jacobs. *Rapid Prototyping and Manufacturing: Fundamentals of Stereo Lithography.* McGraw-Hill, 1993.

[97] G. Jacucci, A. Morrison, G. T. Richard, J. Kleimola, P. Peltonen, L. Parisi, and T. Laitinen. Worlds of information: designing for engagement at a public multitouch display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2267–2276, New York, NY, USA, 2010. ACM.

[98] W. Jaschtnski-Kuza. Visual strain during VDU work: the effect of viewing distance and dark focus. *Ergonomics*, 31(10):1449–1465, 1988. PMID: 3208736.

[99] H. Jiang, E. Ofek, N. Moraveji, and Y. Shi. Direct pointer: direct manipulation for large-display interaction using handheld cameras. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '06, pages 1107–1110, New York, NY, USA, 2006. ACM.

[100] M. Kaltenbrunner, T. Bovermann, B. Ross, and E. Costanza. TUIO - A Protocol for Table Based Tangible User Interfaces. In *Proceedings of the 6th International Workshop on Gesture in Human-Computer Interaction and Simulation*, GW'05, 2005.

[101] N. Kaviani, M. Finke, S. Fels, R. Lea, and H. Wang. What goes where?: designing interactive large public display applications for mobile device interaction. In *Proceedings of the First International Conference on Internet Multimedia Computing and Service*, ICIMCS '09, pages 129–138, New York, NY, USA, 2009. ACM.

[102] C. Kirstein and H. Muller. Interaction with a projection screen using a camera-tracked laser pointer. In *Multimedia Modeling, 1998. MMM '98. Proceedings. 1998*, pages 191–192, 1998.

[103] R. Klanten, S. Ehmann, and V. Hanschke. *A Touch of Code - Interactive Installations and Experiences*. gestalten, 2011.

[104] H. Korhonen and J. Arrasvuori. Let Users Tell the Story: Evaluating User Experience with Experience Prototypes. *Human Factors*, 2010.

[105] H. Korsgaard, N. B. Hansen, D. Basballe, P. Dalsgaard, and K. Halskov. Odenplan: a media façade design process. In *Proceedings of the 4th Media Architecture Biennale Conference: Participation*, MAB '12, pages 23–32, New York, NY, USA, 2012. ACM.

[106] G. Krasner and S. Pope. A cookbook for using the model-view-controller interface paradigm in Smalltalk-80. *Journal of Object Oriented Programming*, 1(3):26–49, 1988.

[107] J.-P. Kreißand G. Neuhaus. *Einführung in die Zeitreihenanalyse*. Springer, 2006.

[108] W. M. Kroner. An intelligent and responsive architecture. *Automation in Construction*, 6(56):381 – 393, 1997.

[109] K. Kuikkaniemi, G. Jacucci, M. Turpeinen, E. Hoggan, and J. Muller. From Space to Stage: How Interactive Screens Will Change Urban Life. *Computer*, 44(6):40 –47, june 2011.

[110] A. Kulik, J. Hochstrate, A. Kunert, and B. Froehlich. The influence of input device characteristics on spatial perception in desktop-based 3D applications. In *3D User Interfaces, 2009. 3DUI 2009. IEEE Symposium on*, pages 59–66, 2009.

[111] C. Lander. *Distanz-Interaktion mit öffentlichen Displays unter Verwendung eines Mobilgeräts*. Saarland University, Saarbrücken, Germany, Department of Computer Science, 2013.

[112] Z. Lei and P. Coulton. A mobile geo-wand enabling gesture based POI search an user generated directional POI photography. In *Proceedings of the International Conference on Advances in Computer Enterntainment Technology*, ACE '09, pages 392–395, New York, NY, USA, 2009. ACM.

[113] M. Lev and J. Snyder. *Map Projections: A Reference Manual*. Taylor & Francis, 1995.

[114] Z. Li, Q. Zhu, and C. Gold. *Digital Terrain Modeling: Principles and Methodology*. CRC Press, 2005.

[115] Y. Lim, E. Stolterman, and J. Tenenberg. The anatomy of prototypes. *ACM Transactions on Computer-Human Interaction*, Vol. 15, ACM, 2008.

[116] A. MacEachren. *How Maps Work*. Guilford Press New York, 1995.

[117] N. Marquardt, R. Diaz-Marino, S. Boring, and S. Greenberg. The proximity toolkit: prototyping proxemic interactions in ubiquitous computing ecologies. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, pages 315–326, New York, NY, USA, 2011. ACM.

[118] P. Marshall, R. Morris, Y. Rogers, S. Kreitmayer, and M. Davies. Rethinking 'multi-user': an in-the-wild study of how groups approach a walk-up-and-use tabletop interface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 3033–3042, New York, NY, USA, 2011. ACM.

[119] A. Maunder, G. Marsden, and R. Harper. Creating and sharing multi-media packages using large situated public displays and mobile phones. In *Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '07, pages 222–225, New York, NY, USA, 2007. ACM.

[120] A. J. Maunder, G. Marsden, and R. Harper. SnapAndGrab: accessing and sharing contextual multi-media content using Bluetooth enabled camera phones and large situated displays. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '08, pages 2319–2324, New York, NY, USA, 2008. ACM.

[121] M. McCullough. *Digital Ground - Architecture, Pervasive Computing, and Environmental Knowing*. The MIT Press, 2005.

[122] D. Mellis, M. Banzi, D. Cuartielles, and T. Igoe. Arduino: an open electronic prototyping platform. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Extended Abstracts*, CHI'07. ACM, 2007.

[123] N. Memarovic, I. Elhart, and M. Langheinrich. Funsquare: first experiences with autopoiesic content. In *Proceedings of the 10th International Conference on Mobile and Ubiquitous Multimedia*, MUM '11, pages 175–184, New York, NY, USA, 2011. ACM.

[124] N. Memarovic, V. Kostakos, G. Fitzpatrick, and A. Schmidt. SOFTec 2013: second workshop on computer mediated social offline interactions. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, pages 1373–1380, New York, NY, USA, 2013. ACM.

[125] N. Memarovic, M. Langheinrich, F. Alt, I. Elhart, S. Hosio, and E. Rubegni. Using public displays to stimulate passive engagement, active engagement, and discovery in public spaces. In *Proceedings of the 4th Media Architecture Biennale Conference: Participation*, MAB '12, pages 55–64, New York, NY, USA, 2012. ACM.

[126] D. Michelis and J. Müller. The Audience Funnel: Observations of Gesture Based Interaction With Multiple Large Displays in a City Center. *International Journal of Human-Computer Interaction*, 27(6):562–579, 2011.

[127] L. Mignonneau and C. Sommerer. Media facades as architectural interfaces. In C. Sommerer, L. Jain, and L. Mignonneau, editors, *The Art and Science of Interface and Interaction Design*, volume 141 of *Studies in Computational Intelligence*, pages 93–104. Springer Berlin Heidelberg, 2008.

[128] A. V. Moere and D. Hill. Research through design in the context of teaching urban computing. In *Street Computing Workshop (co-located with OZCHI09)(Melbourne, Novermber 23–27, 2009)*, 2009.

[129] A. V. Moere and N. Wouters. The role of context in media architecture. In *Proceedings of the 2012 International Symposium on Pervasive Displays*, PerDis '12, pages 12:1–12:6, New York, NY, USA, 2012. ACM.

[130] B. Moggridge. *Designing Interactions*. MIT Press, 2006.

[131] J. Müller, F. Alt, D. Michelis, and A. Schmidt. Requirements and design space for interactive public displays. In *Proceedings of the International Conference on Multimedia*, MM '10, pages 1285–1294, New York, NY, USA, 2010. ACM.

[132] B. Myers, C. Peck, J. Nichols, D. Kong, and R. Miller. Interacting at a distance using semantic snarfing. In G. Abowd, B. Brumitt, and S. Shafer, editors, *Ubicomp 2001: Ubiquitous Computing*, volume 2201 of *Lecture Notes in Computer Science*, pages 305–314. Springer Berlin Heidelberg, 2001.

[133] B. A. Myers, R. Bhatnagar, J. Nichols, C. H. Peck, D. Kong, R. Miller, and A. C. Long. Interacting at a distance: measuring the performance of laser pointers and other devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '02, pages 33–40, New York, NY, USA, 2002. ACM.

[134] Y. Nakanishi. Virtual Prototyping Using Miniature Model and Visualization for Interactive Public Displays. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, pages 458–467, New York, NY, USA, 2012. ACM.

[135] Y. Nakanishi, K. Sekiguchi, T. Ohmori, S. kitahara, and D. Akatsuka. Hybrid Prototyping by Using Virtual and Miniature Simulation for Designing Spatial Interactive Information Systems. In K. Lyons, J. Hightower, and E. Huang, editors, *Pervasive Computing*, volume 6696 of *Lecture Notes in Computer Science*, pages 250–257. Springer Berlin Heidelberg, 2011.

[136] J. Nielsen. *Usability Engineering*. Morgan Kaufmann, 1994.

[137] K. O'Hara, M. Perry, E. Churchill, and D. Russel, editors. *Public and Situated Displays - Social Interactional Aspects of Shared Display Technologies*. Kluwer Academic Publishers, 2003.

[138] T. Ojala, H. Kukka, T. Lindn, T. Heikkinen, M. Jurmu, S. Hosio, and F. Kruger. UBI-Hotspot 1.0: large-scale long-term deployment of interactive public pisplays in a city center. In *Internet and Web Applications and Services (ICIW), 2010 Fifth International Conference on*, pages 285–294, 2010.

[139] D. R. Olsen, Jr. and T. Nielsen. Laser pointer interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, pages 17–22, New York, NY, USA, 2001. ACM.

[140] T. Paek, M. Agrawala, S. Basu, S. Drucker, T. Kristjansson, R. Logan, K. Toyama, and A. Wilson. Toward universal mobile interaction for shared displays. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, CSCW '04, pages 266–269, New York, NY, USA, 2004. ACM.

[141] M. Pasupathi, K. McLean, and T. Weeks. To Tell or Not to Tell: Disclosure and the Narrative Self. *Journal of Personality*, 77, 2009.

[142] E. Paulos and E. Goodman. The familiar stranger: anxiety, comfort, and play in public places. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pages 223–230, New York, NY, USA, 2004. ACM.

[143] E. Paulos, R. Honicky, and B. Hooker. Citizen science: enabling participatory urbanism. *Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City*, pages 414–436, 2009.

[144] N. Pears, D. Jackson, and P. Olivier. Smart phone interaction with registered displays. *IEEE Pervasive Computing*, 8(2):14–21, 2009.

[145] P. Peltonen, E. Kurvinen, A. Salovaara, G. Jacucci, T. Ilmonen, J. Evans, A. Oulasvirta, and P. Saarikko. It's mine, don't touch!: interactions at a large multi-touch display in a city centre. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1285–1294, New York, NY, USA, 2008. ACM.

[146] P. Peltonen, A. Salovaara, G. Jacucci, T. Ilmonen, C. Ardito, P. Saarikko, and V. Batra. Extending large-scale event participation with user-created mobile media on a public display. In *Proceedings of the 6th International Conference on Mobile and Ubiquitous Multimedia*, MUM '07, pages 131–138, New York, NY, USA, 2007. ACM.

[147] B. Phong. Illumination for computer generated pictures. *Commmunications of the ACM*, 18(6):311–31, 1975.

[148] S. Reeves, S. Benford, C. O'Malley, and M. Fraser. Designing the spectator experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, pages 741–750, New York, NY, USA, 2005. ACM.

[149] D. Reilly, M. Welsman-Dinelle, C. Bate, and K. Inkpen. Just point and click?: using handhelds to interact with paper maps. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services*, MobileHCI '05, pages 239–242, New York, NY, USA, 2005. ACM.

[150] J. Rekimoto. Tilting operations for small screen interfaces. In *Proceedings of the 9th Annual ACM Aymposium on User Interface Software and Technology*, UIST '96, pages 167–168, New York, NY, USA, 1996. ACM.

[151] J. Rekimoto and M. Saitoh. Augmented surfaces: a spatially continuous work space for hybrid computing environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pages 378–385, New York, NY, USA, 1999. ACM.

[152] T. Reynolds and J. Gutman. Laddering Theory, Method, Analysis, and Interpretation. *Journal of Advertising Research*, Vol. 28, 1988.

[153] J. Rudd, K. Stern, and S. Isensee. Low vs. high-fidelity prototyping debate. *ACM Interactions*, Vol. 3, ACM, 1996.

[154] E. Rukzio, K. Leichtenstern, V. Callaghan, P. Holleis, A. Schmidt, and J. Chin. An experimental comparison of physical mobile interaction techniques: touching, pointing and scanning. In P. Dourish and A. Friday, editors, *UbiComp 2006: Ubiquitous Computing*, volume 4206 of *Lecture Notes in Computer Science*, pages 87–104. Springer Berlin Heidelberg, 2006.

[155] D. Saffer. *Designing for Interaction.* New Riders, 2006.

[156] L. Sanchez, V. Gutierrez, J. A. Galache, P. Sotres, J. R. Santana, J. Casanueva, and L. Munoz. SmartSantander: experimentation and service provision in the smart city. In *16th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, pages 1–6, 2013.

[157] J. Scheible and T. Ojala. MobiSpray: mobile phone as virtual spray can for oainting big anytime anywhere on anything. *Leonardo*, 42, No. 4:332 – 341, 2009.

[158] J. Scheible, T. Ojala, and P. Coulton. MobiToss: a novel gesture based interface for creating and sharing mobile multimedia art on large public displays. In *Proceedings of the 16th ACM International Conference on Multimedia*, MM '08, pages 957–960, New York, NY, USA, 2008. ACM.

[159] D. Schmidt, F. Chehimi, E. Rukzio, and H. Gellersen. PhoneTouch: a technique for direct phone interaction on surfaces. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, pages 13–16, New York, NY, USA, 2010. ACM.

[160] D. Schmidt, J. Seifert, E. Rukzio, and H. Gellersen. A cross-device interaction style for mobiles and surfaces. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, pages 318–327, New York, NY, USA, 2012. ACM.

[161] O. Schoch. My building is my display. In *Proc. eCAADe'06*, 2006.

[162] S. D. Scott, M. S. T. Carpendale, and K. M. Inkpen. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, CSCW '04, pages 294–303, New York, NY, USA, 2004. ACM.

[163] S. Seitinger, D. S. Perry, and W. J. Mitchell. Urban pixels: painting the city with light. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 839–848, New York, NY, USA, 2009. ACM.

[164] J. Serra and L. Vincent. An overview of morphological filtering. *Circuits, Systems and Signal Processing*, 11(1):47–108, 1992.

[165] D. A. Shamma, J. Scheible, and R. M. Sheppard. Graffiti dance: interaction of light, information, and environment. In *Proceedings of the 7th ACM Conference on Creativity and Cognition*, C&C '09, pages 479–480, New York, NY, USA, 2009. ACM.

[166] H. Sharp, Y. Rogers, and J. Preece. *Interaction Design - Beyond Human-Computer Interaction*. John Wiley & Sons,, 2011.

[167] K. Sheldon, A. Elliot, Y. Kim, and T. Kasser. What is satisfying about satisfying events? *Journal of Personality and Social Psychology*, 2001.

[168] G. Shoemaker, A. Tang, and K. S. Booth. Shadow reaching: a new perspective on interaction for large displays. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology*, UIST '07, pages 53–56, New York, NY, USA, 2007. ACM.

[169] R. Simon and P. Fröhlich. A mobile application framework for the geospatial web. In *Proceedings of the 16th International Conference on World Wide Web*, WWW '07, pages 381–390, New York, NY, USA, 2007. ACM.

[170] A. Skupin. From metaphor to method: cartographic perspectives on information visualization. In *Information Visualization, 2000. InfoVis 2000. IEEE Symposium on*, pages 91–97, 2000.

[171] J. Snyder, editor. *Flattening the Earth - Two Thousand Years of Map Projections*. The University of Chicago Press, 1993.

[172] C. Stahl. *Spatial modeling of activity and user assistance in instrumented environments*. PhD thesis, Saarländische Universitäts- und Landesbibliothek, Postfach 151141, 66041 Saarbrücken, 2009.

[173] C. Stahl and J. Haupert. Simulating and evaluating public situated displays in virtual environment models. In *Proceedings of MODIE'06*. ACM, 2006.

[174] C. Stahl and J. Haupert. Taking location modelling to new levels: a map modelling toolkit for intelligent environments. In M. Hazas, T. Strang, and J. Krumm, editors,

Proceedings of the International Workshop on Location- and Context-Awareness (LoCA), LNCS 3987, volume 3987 / 2006 of *Lecture Notes in Computer Science*, pages 74–85, Munich, Germany, 2006. Springer-Verlag Berlin Heidelberg.

[175] U. Stalder. Digital Out-of-Home Media: Means and Effects of Digital Media in Public Space. In J. Müller, F. Alt, and D. Michelis, editors, *Pervasive Advertising*, Human-Computer Interaction Series, pages 31–56. Springer London, 2011.

[176] R. Stoakley, M. J. Conway, and R. Pausch. Virtual Reality on a WIM: Interactive Worlds in Miniature. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '95, pages 265–272, New York, NY, USA, 1995. ACM Press/Addison-Wesley Publishing Co.

[177] M. Struppek. Urban screens - the urbane potential of public screens for interaction. *Visual Communication*, 5:173 – 188, 2006.

[178] R. Su and B. Bailey. Put them where? towards guidelines for positioning large displays in interactive workspaces. In M. Costabile and F. Patern, editors, *Human-Computer Interaction - INTERACT 2005*, volume 3585 of *Lecture Notes in Computer Science*, pages 337–349. Springer Berlin Heidelberg, 2005.

[179] M. Tani, K. Yamaashi, K. Tanikoshi, M. Futakawa, and S. Tanifuji. Object-oriented video: interaction with real-world objects through live video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '92, pages 593–598, New York, NY, USA, 1992. ACM.

[180] A. M. Townsend. *Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia*. W. W. Norton & Company, 2013.

[181] E. Toye, R. Sharp, A. Madhavapeddy, D. Scott, E. Upton, and A. Blackwell. Interacting with mobile services: an evaluation of camera-phones and visual tags. *Personal and Ubiquitous Computing*, 11(2):97–106, 2007.

[182] G. Tscherteu. *The Media Architecture Compendium*. mediaarchitecture.org, 2013.

[183] J. Tyner. *Principles of Map Design*. The Guilford Press, New York, 2010.

[184] United Nations, New York. *World Urbanization Prospects: The 2007 Revision*. Department of Economic and Social Affairs, Population Division, 2007.

[185] United Nations, New York. *World Urbanization Prospects: The 2011 Revision*. Department of Economic and Social Affairs, Population Division, 2011.

[186] T. Vajk, P. Coulton, W. Bamford, and R. Edwards. Using a mobile phone as a "Wii-like" controller for playing games on a large public display. *International Journal of Computer Games Technology*, 2008:173 – 188, 2008.

[187] D. Vogel and R. Balakrishnan. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology*, UIST '04, pages 137–146, New York, NY, USA, 2004. ACM.

[188] F. Vogt, J. Wong, B. Po, R. Argue, S. Fels, and K. Booth. Exploring collaboration with group pointer interaction. In *Computer Graphics International, 2004. Proceedings*, pages 636–639, 2004.

[189] R. Want, K. P. Fishkin, A. Gujar, and B. L. Harrison. Bridging physical and virtual worlds with electronic tags. In *CHI*, volume 99, pages 370–377, 1999.

[190] D. Watson, L. Clark, and A. Tellegen. Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, 54, 1988.

[191] M. Weiser. The computer for the 21st century. *Scientific American*, 265(3):94–104, 1991.

[192] A. Wiethoff. *Prototyping Tools for Hybrid Interactions*. PhD thesis, Ludwig-Maximilians-Universität München, Dezember 2012.

[193] A. Wiethoff and M. Blöckner. LightBox: Exploring Interaction Modalities with Colored Light. In *Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '11, pages 399–400, New York, NY, USA, 2011. ACM.

[194] A. Wiethoff and S. Gehring. Designing interaction with media façades: a case study. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, pages 308–317, New York, NY, USA, 2012. ACM.

[195] A. Wiethoff and S. Gehring. Interacting with light. In R. Wichert, K. Laerhoven, and J. Gelissen, editors, *Constructing Ambient Intelligence*, volume 277 of *Communications in Computer and Information Science*, pages 63–67. Springer Berlin Heidelberg, 2012.

[196] K. S. Willis, G. Roussos, K. Chorianopoulos, and M. Struppek. Shared encounters. In K. S. Willis, G. Roussos, K. Chorianopoulos, and M. Struppek, editors, *Shared Encounters*, 1431-1496, pages 1–15, London, UK, 2020. Springer.

[197] A. Wilson and S. Shafer. XWand: UI for intelligent spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '03, pages 545–552, New York, NY, USA, 2003. ACM.

[198] A. D. Wilson and H. Benko. Combining multiple depth cameras and projectors for interactions on, above and between surfaces. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, pages 273–282, New York, NY, USA, 2010. ACM.

[199] C. Wingrave, Y. Haciahmetoglu, and D. Bowman. Overcoming world in miniature limitations by a scaled and scrolling wim. In *Proceedings of 3DUI'06*, pages 11–16. IEEE, 2006.

[200] C. Winkler, K. Pfeuffer, and E. Rukzio. Investigating mid-air pointing interaction for projector phones. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*, ITS '12, pages 85–94, New York, NY, USA, 2012. ACM.

[201] T. Winograd. *Interaction Spaces for Twenty-First-Century Computing*, pages 259–276. Addison-Wesley, 2001.

[202] P. Wright and J. McCarthy. Empathy and Experience in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 637–646, New York, NY, USA, 2008. ACM.

[203] P. Wright and J. McCarthy. Experience-centered design: Designers, users and communities in dialogue. *Synthesis Lectures on Human-Centered Informatics*, 3, 2010.

[204] K. Yatani, K. Tamura, K. Hiroki, M. Sugimoto, and H. Hashizume. Toss-it: intuitive information transfer techniques for mobile devices. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1881–1884, New York, NY, USA, 2005. ACM.

[205] Y. Zhang, A. Bulling, and H. Gellersen. SideWays: a gaze interface for spontaneous interaction with situated displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 851–860, New York, NY, USA, 2013. ACM.