

Multi-User Interaction on Media Facades through Live Video on Mobile Devices

Sebastian Boring¹, Sven Gehring², Alexander Wiethoff¹,
Magdalena Blöckner¹, Johannes Schöning², Andreas Butz¹

¹University of Munich, Munich, Germany

{sebastian.boring, alexander.wiethoff, andreas.butz}@ifi.lmu.de, bloeckner@cip.ifi.lmu.de

²German Research Center for Artificial Intelligence, Saarbrücken, Germany
{sven.gehring, johannes.schoening}@dfki.de

ABSTRACT

The increasing number of media facades in urban spaces offers great potential for new forms of interaction – especially for collaborative multi-user scenarios. In this paper, we present a way to directly interact with them *through live video* on mobile devices. We extend the *Touch Projector* interface to accommodate multiple users by showing individual content on the mobile display that would otherwise clutter the facade’s canvas or distract other users. To demonstrate our concept, we built two collaborative multi-user applications: (1) painting on the facade and (2) solving a 15-puzzle. We gathered informal feedback during the ARS Electronica Festival in Linz, Austria and found that our interaction technique is (1) considered easy-to-learn, but (2) may leave users unaware of the actions of others.

Author Keywords

Mobile device, media facades, interaction techniques, input device, interface distribution, augmented reality.

ACM Classification Keywords

H5.2 [Information interfaces and presentation]: User Interfaces – Input devices and strategies, Interaction styles, Graphical user interfaces.

General Terms

Experimentation, Human Factors.

INTRODUCTION

More and more urban landscapes are equipped with media facades. The Beijing National Aquatics Center in Beijing, China and the ARS Electronica center in Linz, Austria are two prominent examples out of hundreds of such facades. However, due to their size and the required viewing distance, interacting with them directly (e.g., by touching them) is normally impossible. Recent advances in mobile

computing allow users to interact with such facades in several ways. Current interaction approaches include controlling pointers on the facade’s canvas [2] or pushing content to it through multimedia messages [8].

The size, visibility, and large audience of media facades offer a great potential for collaborative interaction. Indirect pointing techniques, however, restrict the number of simultaneous users: (1) each pointer occludes a (small) portion of the facade, eventually leading to clutter and large content regions being invisible. (2) More pointers shown on the facade make finding one’s own pointer rather difficult. In addition, facades showing pointers need reasonably high resolutions to provide enough pixels per pointer. One approach to solve these issues is to use an absolute and direct technique such as interaction *through live video* [3].

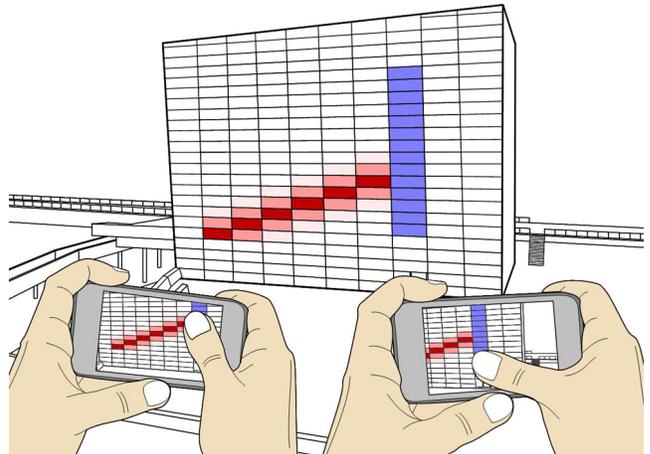


Figure 1. Interacting through live video allows multiple users to manipulate a media facade. Changes (also those of other users) are shown immediately on the facade and the mobile device. Colors denote actions from other users.

In this paper, we apply and extend the concept of *Touch Projector* [3] to media facades to avoid the limitations of current interaction techniques. With two applications, we demonstrate how multiple users can interact on a facade simultaneously. During the ARS Electronica Festival we found that our approach is (1) considered easy-to-learn, but (2) may leave users unaware of the actions of others.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.

Copyright 2011 ACM 978-1-4503-0267-8/11/05...\$10.00.

RELATED WORK

The term “media facade” describes the idea of designing or modifying the architecture of buildings with the objective of using their surfaces as giant public screens [6,13]. In addition, urban public spaces are emerging more and more as prime locations for media facades that are embedded in the landscape of a city [11]. Researchers recently explored the social potential of such media facades as they can be seen or even designed by multiple persons simultaneously [10,12]. They further studied public participatory crowd experiences when porting popular games to media facades in combination with mobile devices [4]. Dalsgaard et al. described eight key challenges when designing such novel interactive systems and – most importantly – when offering users new, distributed interfaces [5].

Several techniques have been proposed to interact with distant displays. The most prominent techniques are relative and indirect pointing as well as augmented reality approaches. Relative and indirect pointing can be used to distant displays by turning a camera-equipped mobile device into a mouse-like device [1]. However, such input techniques may hinder multi-user scenarios due to the required virtual pointer. *MobiSpray* uses a world-in-miniature interface to allow for “spraying” color on various projected surfaces using the mobile device’s accelerometers [9]. Recent advantages in mobile augmented reality allow absolute pointing on displays [7]. For tracking purposes, their system relies on a marker shown on the remote display. *Touch Projector* allows interaction with a distant display shown in the viewfinder using touch in real-time without relying on fiducial markers [3]. As this system follows a direct input approach, we decided to take it as the basis of our prototype.

DESIGNING INTERACTION ON MEDIA FACADES

Our goal was to implement a system that allows multiple users to interact simultaneously on a media facade. As discussed in the previous section, relative and/or indirect approaches (e.g., *Sweep* [1]) may limit the number of users to the number of distinguishable (i.e., identifiable) pointers on the remote canvas. The low resolution facade of the ARS Electronica Center¹ (each window or pixel is about one by three meters in size) used for our prototypes lowers this number even more. Techniques that use a world-in-miniature representation (e.g., *MobiSpray* [9]) overcome this limitation at the expense of macro attention shifts between both the mobile and target display (here: facade).

To avoid the necessity of virtual pointers as well as the potential costs of macro attention shifts, we decided to use the concept of *Touch Projector* on media facades [3]. Users aim their device at the facade and observe it in live video, allowing them to point *through the display*. Touch input occurring on the mobile device is *projected* onto the facade,

giving the impression that users directly touch the building (see Figure 1). Using this concept on large media facades in the wild, however, comes with several challenges.

Technical challenges of interacting with media facades

According to Dalsgaard et al., applications must consider potential shifts in lighting and weather conditions [5]. The facade of the ARS Electronica Center is only visible below a certain level of daylight. The original implementation of *Touch Projector* did not account for this, as it was built for regular computer screens with strong background lighting in controlled environments. Especially in the dark, reflections on wet ground are commonly caused by weather. As described in the IMPLEMENTATION section, we substantially changed the tracking algorithm to allow for outdoor use.

Another challenge is that media facades mostly have unique features (e.g., different sizes, resolutions, and optimal viewing distances). The ARS Electronica Center can easily be viewed at a distance of 300 meters. However, the distance influences the facade’s apparent size in live video. To counter this, we used the zoom functionality of *Touch Projector*. Ideally, the building would fit exactly into the live video image. If this is not the case, the mobile device adjusts its zoom level. This ensures a practically constant control-display ratio for users independently of their distance.

Allowing multi-user interaction on media facades

The large size of such facades further allows multiple users to interact simultaneously on them. The original idea of *Touch Projector* only transforms input occurring on the mobile device to a facade’s canvas. Thus, interactive controls or temporary feedback are shown on the facade. This is not always an optimal solution: (1) tool palettes waste screen real estate decreasing the size of the actual interaction canvas. (2) The resolution of the facade (in combination with the viewing distance) further limits the resolution of such controls. (3) Temporary feedback (e.g., highlighted regions) interferes with the interaction of others.

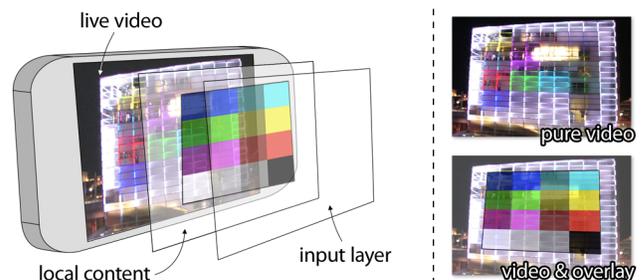


Figure 2. Inserting a local content layer (left) allows for showing pure or augmented live video for each user individually.

On the other hand, the live video on the mobile device shows the facade at all times. Thus, the facade’s content can be augmented locally without introducing macro attention shifts. To allow for such feedback, the mobile screen superimposes a personal layer on the local live video (see Figure 2), leaving the shared view of the facade canvas unaffected.

¹ The facade of the ARS Electronica Center hosts about 40000 LEDs embedded into 1087 addressable windows.

IMPLEMENTATION AND DEPLOYMENT

Our prototype uses a dedicated server to (1) control the building's lighting through DMX and (2) communicate with mobile devices (here: three Apple iPhones). Similar to *Touch Projector*, the mobile devices send video frames over wireless LAN to this server, which calculates their spatial relationship to the building. The server further handles all touch events received from the mobile devices.

Addressing tracking challenges for outdoor use

To identify the facade, we chose to show a white frame around an entire side of the building by permanently lighting the outmost pixels. This frame can be detected using *Touch Projector's* image processing methods (i.e., contrast correction, edge detection, and corner identification). Our system then uses the detected perspective distortion of the building's outline (i.e., the frame), to calculate the spatial relationship between mobile device and building.

To avoid reflections on wet surfaces being falsely detected, we made two assumptions: (1) users point their devices at the building instead of the reflection, which causes the reflection to be shown only partly. (2) The reflection is slightly jittered resulting in less prominent lines of the building's outline. In early tests on the facade with real users and reflections caused by wet ground around the building, we found these assumptions to be sufficient.

Allowing for personal content on the mobile device

Aside from determining the spatial relationship of mobile devices, the server stores individual content for each user as image data which (on request) can be transferred to the mobile device. In some cases (i.e., directly superimposing individual content), the system also distorts the content for correct alignment with live video. This is done by using the inverted transformation matrix (i.e., homography) calculated during the detection process. Once the image is sent to the mobile device, it is overlaid on the live video image.



Figure 3. On request, the original video image of the puzzle (a) is augmented with a grid (b) or a preview (c).

All interaction events (i.e., touch inputs) are sent to the server regardless whether the user hit a “local” item or not. As the server knows the exact locations of all elements it can determine and execute the associated action. Thus, application developers only need to design the interface elements and their actions on the server. This type of implementation allows for greater flexibility in terms of the heterogeneity found in mobile device platforms. It limits scalability, however, since computation on the server linearly increases with the number of mobile devices.

Example applications at ARS Electronica

To demonstrate the use of (1) interacting through live video on media facades as well as (2) the distribution of public and personal content, we built two applications. These allowed users to paint freely on the facade or solve a 15-puzzle in a collaborative way.

Our first application allows users to solve a 15-puzzle on the facade. Eight pixels (i.e., 2 by 4 windows) are representing one tile of the puzzle. Tiles can be shifted by tapping on a tile next to the missing one. However, the facade's low resolution did not allow for clearly showing division lines that are important to identify tiles (see Figure 3a). We decided to allow users to superimpose these lines on the mobile device to allow for tile identification (see Figure 3b). As our tiles only have 8 pixels in total, it is hard to identify a tile's correct location. Users can peek at the solved puzzle by requesting a preview (see Figure 3c). We decided to show the preview locally so that others are not distracted.



Figure 4. To keep the drawing canvas as large as possible (a), users can switch to the tool palette (b).

Our second application allows users to paint freely on the facade. Similar to common drawing applications, users (1) choose a color and (2) select a tool from a tool palette. To do so, users perform a slide gesture next to the live video image. The mobile device then shows a tool palette (see Figure 4b). After closing the palette (i.e., sliding in opposite direction), users can apply the selected color and tool to the building by touching (and dragging) on it in live video (see Figure 4a). Placing the controls on the mobile device was the only possible solution, as our facade does not offer a resolution high enough to display controls.

Initial user feedback

During the ARS Electronica Festival in Linz, Austria, we presented our applications to a broad audience. We handed phones with the application already running to users without any further instructions. By observing how others used the application, they immediately started to interact with the facade. Up to three persons were able to interact simultaneously, but we ensured that at least two did at all times. Downloading the application was not possible as (1) we used a restricted network and (2) it was not allowed in the AppStore at that time. Nevertheless, with three users interacting simultaneously, we were able to observe interesting scenarios including collaboration between them.

Out of the approximately 50 users we asked 15 attendees (5 female; average age was 26.1) for feedback after interacting with the building. In informal interviews we found that this style of interaction is perceived as (1) easy-to-learn and (2)

easy-to-use. Overall, the feedback we gained during the interviews was highly positive. The fact that they could directly change the facade in real time (e.g., a form of *digital graffiti*) was mentioned positively. However, users were sometimes annoyed by the parallel use of our application. The most important statement was: “It is good to interact in a parallel way if you know the person. But if you don’t know the person, you are kind of fighting over pixels and space to draw. It’s kind of annoying”. While this user favored collaboration, another pair of users created a strobe-like effect, alternately filling the entire facade with white and black. Thus, interactions involving either collaboration or competition were supported by the painting application.

CONCLUSIONS & FUTURE WORK

In this paper, we presented an extension to the concept of *Touch Projector* to allow multiple users to interact collaboratively (or competitively) with media facades shown in live video on their mobile device. We described the technical realization that can be used under various weather conditions on any digital surface that has or can display a white frame. We further extended *Touch Projector* by superimposing individual content (i.e., UI elements that are not of interest for all users at a time) on the live video. While this was necessary for the low resolution facade in our deployment, it constitutes a very general mechanism when many users 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.

We demonstrated our prototype during the ARS Electronica Festival in Linz, Austria with a large group of users. The feedback we gained informs future work in the area of multi-user interaction at-a-distance. In contrast to collocated scenarios in which users are next to or can see one another, larger facades may give rise to greater distances between users, so that they may not be aware of (1) *who* is interacting and (2) *where* others are. As this is a common problem of techniques that use interaction at-a-distance, we plan to develop solutions to the awareness problem, by, for example, visualizing the location and direction of others.

Another issue raised by our participants was the heavily parallel nature of interaction using our technique. The fact that users could simultaneously interact in the same region of the facade was only appreciated if users knew each other. Otherwise, they rather got frustrated if others interacted (and thus interfered) with them in “their” region. There will always be some tension between permitting desired interactions and preventing undesired ones on a large-scale, multi-user, public media facade. As this is an intrinsic property of the medium and not solvable in general, we hope to iteratively converge on a more appropriate balance, through partitioning time slots or sub-regions among users on the facade with the ultimate goal of maximizing enjoyment and minimizing frustrations for future users.

ACKNOWLEDGMENTS

This work has been funded by both the “Deutsche Forschungsgemeinschaft” (DFG) and the German State of Bavaria. We thank Antonio Krüger and Michael Rohs for their input in the initial design phase. We also thank Andreas Pramböck, Stefan Mittelböck and Horst Hörtnner (AEC) for their technical support during preparation as well as the festival. We further thank Dominikus Baur and – especially – Joe McCarthy for their valuable comments and feedback.

REFERENCES

1. Ballagas, R., Borchers, J., Rohs, M., Sheridan, J.G. The Smart Phone: A Ubiquitous Input Device. *IEEE Pervasive Computing* 5, 1 (2006), 70-77.
2. Boring, S., Jurmu, M., Butz, A. Scroll, tilt or move it: using mobile phones to continuously control pointers on large public displays. *Proc. OzCHI 2009*, ACM Press (2009), 161-168.
3. Boring, S., Baur, D., Butz, A., Gustafson, S., Baudisch, P. Touch projector: mobile interaction through video. *Proc. CHI 2010*, ACM Press (2010), 2287-2296.
4. Borries, F., Walz, S., Böttger, M. Space Time Play: Computer Games, Architecture and Urbanism: The Next Level, *Birkhäuser, Basel*, 2007, 396-397.
5. Dalsgaard, P., Halskov, K. Designing Urban Media Facades: Cases and Challenges. *Proc. CHI '10*, ACM Press (2010), 2277-2286.
6. Häusler M.H. Media Facades: History, Technology, Content. *Av Edition*, Ludwigsburg, 2009.
7. Pears, N., Jackson, D.G., Olivier, P. Smart Phone Interaction with Registered Displays. *IEEE Pervasive Computing* 8, 2 (2009), 14-21.
8. Peltonen, P., Salovaara, A., Jacucci, G., Ilmonen, T., Ardito, C., Saarikko, P., Batra, V. Extending large-scale event participation with user-created mobile media on a public display. *Proc. MUM '07*, ACM (2007), 131-138.
9. Scheible, J., Ojala, T. MobiSpray: mobile phone as virtual spray can for painting BIG anytime anywhere on anything. *Proc. SIGGRAPH '09*, ACM Press, 332-341.
10. Schoch, O. My Building is my Display, *Proc. eCAADe '06*, 2006, 610-616.
11. Seiting, S., Perry, D.S., Mitchell, W.J., Urban Pixels: Painting the City with Light. *Proc. CHI 2009*, 839-848
12. Sommerer, C., Jain, L., Mignonneau, L. Media Facades as Architectural Interfaces. *The Art and Science of Interface and Interaction Design*, Springer, 2008, 93-104
13. Struppek, M. The social potential of Urban Screens, *Visual Communication* 5, 2, 2006; 173-188.
14. Tani, M., Yamashi, K., Tanikoshi, K., Futakawa, M., Tanifuji, S. Object-oriented video: interaction with real-world objects through live video. *Proc. CHI 1992*, ACM Press (1992), 593-598.