Combining touch & mid-air gestures on public displays

Thor Bossuyt and Niklas Hillgren

LIU-IDA/LITH-EX-A--12/068--SE

2012-11-07
Examensarbete

Combining touch & mid-air gestures on public displays

av

Thor Bossuyt and Niklas Hillgren

LIU-IDA/LITH-EX-A--12/068--SE

2012-11-07

Examinator: Erik Berglund
Handledare: Anders Fröberg
This thesis investigates the impact of different factors on public displays that use touch and mid-air gestures. We present a novel application for public displays: MirrorTouch — a game combining touch interaction with mid-air gesturing through depth sense. We studied the impact of these factors through a series of field studies. We show that clear affordances are important to communicate both modalities clearly and also how to improve the conversion rate of passers-by. We present our findings of how social situations around such a display varied with the location where the display was placed. We also present quantitative data on group behavior around our public display, such as a measuring of the honeypot effect and how people made transitions between the different modalities. Our results can be of great value to designers and researchers of public displays that want to deploy displays with similar techniques.
ACKNOWLEDGMENTS

We would like to thank the employees that work at Quality and Usability Lab of Deutsche Telekom Laboratories. The warm and loving working environment created a superb atmosphere, both of us felt welcome there since day one.

A special thanks goes out to the members of our team consisting of Jörg Müller, Gilles Bailly, Robert Walter, Constantin Schmidt and Dieter Eberle.

Last but not least we would like to thank Linköping University and our examiner Erik Berglund. Without you it wouldn’t be possible for us to go abroad.
# CONTENTS

## I INTRODUCTION

### 1 BACKGROUND

- 1.1 Public displays
- 1.2 Problem
  - 1.2.1 Problem statement
  - 1.2.2 Subproblems
  - 1.2.3 Limitations

### 1.3 Outline

## II RELATED WORK

- 2.1 Combined mid-air interaction with touch
- 2.2 Field studies
- 2.3 Stages of interaction
- 2.4 Design constraints
- 2.5 Live mic problem
- 2.6 The Honeypot Effect

## III METHODS

- 3.1 Field studies
  - 3.1.1 Application
  - 3.1.2 Locations
  - 3.1.3 Logging and observing

- 3.2 Statistical validity
  - 3.2.1 Inter-rater reliability
  - 3.2.2 Chi-square

## II RESULTS

## IV THEORY STUDY: MODALITIES

- 4.1 Touch and mid-air gestures
- 4.2 State-transition
  - 4.2.1 Primitives
  - 4.2.2 Displacement and Clicking
  - 4.2.3 Touch and mid-air objects
  - 4.2.4 Our application

## V FIELD STUDY: LOCATION STUDY

## VI FIELD STUDY: NOTICING TOUCH INTERACTION

- 6.1 Background & Goal
- 6.2 Hypothesis
- 6.3 Setup
- 6.4 Results
- 6.5 Discussion and Validity

## VII FIELD STUDY: QUANTIFYING THE HONEYPOT EFFECT AND THE N/T POINTS
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>Background &amp; Goal</td>
<td>31</td>
</tr>
<tr>
<td>7.2</td>
<td>Hypothesis</td>
<td>31</td>
</tr>
<tr>
<td>7.3</td>
<td>Setup</td>
<td>31</td>
</tr>
<tr>
<td>7.4</td>
<td>Results</td>
<td>32</td>
</tr>
<tr>
<td>7.4.1</td>
<td>Positioning</td>
<td>32</td>
</tr>
<tr>
<td>7.4.2</td>
<td>Quantification of honeypot effect</td>
<td>33</td>
</tr>
<tr>
<td>7.5</td>
<td>Discussion and Validity</td>
<td>34</td>
</tr>
</tbody>
</table>

### III Conclusion and Discussion 37

8 Conclusion 39

9 Discussion 41

9.1 Reflection 41

9.2 Future Work 41

### IV Appendix 43

A MirrorTouch Locations Study: A Research Paper Submitted to CHI’13 45

Bibliography 51
LIST OF FIGURES

Figure 1 Different types of content in the application. 12
Figure 2 Research corridor 14
Figure 3 University corridor 15
Figure 4 University hall 15
Figure 5 Exhibition hall 15
Figure 6 Two of the researchers observing from an inconspicuous location. 16
Figure 7 State-transition model of input 19
Figure 8 Clutching for mid-air gestures 20
Figure 9 Interaction primitives for input. Left: cursor input, right: shape input technique 21
Figure 10 When the mirror image displaces objects, it is impossible to touch them. 21
Figure 11 Heatmap of all passers-by, university corridor (as seen from above) 32
Figure 12 Heatmap of people touching, university corridor (as seen from above) 33
Figure 13 Heatmap of people mid-air gesturing, university corridor (as seen from above) 33
Figure 14 Conversion diagram with different preconditions. The first plot is the baseline where all passers-by during three days are measured. The last three measure interactions within the pre-defined honeypot timeframe for both modalites, for touch alone and for mid-air gesturing alone. Based on 1701 passers-by. 34
Figure 15 Histogram over the group sizes that were seen interacting during a two day observation of the university corridor. Based on 81 interaction sessions. 34

LIST OF TABLES

Table 1 Design constraints fulfilled 13
Table 2 Summary of the different locations. 14
Using either clicking or displacement for both modalities mirrors the same functionality for mid-air and touch. The utility of combining touch displacement with distant click is unclear. Combining distant displacement with touch click raises potential problems (see Figure 10).

When mid-air objects influence touch objects, the conflict shown in figure 10 can occur. If touch objects influence mid-air objects, but not vice versa, content integrates both modalities. If neither objects can influence another, two entirely separate contents (mid-air and touch) are shown on the same display.

Experiment conditions of field study 2

Day 1: Passers-by touch conversion rate

Day 2: Passers-by touch conversion rate

Day 1 & 2 combined, passers-by touch conversion rate

Day 1 & 2 combined, passers-by interaction conversion rate

ACRONYMS

DOF  Degrees of freedom

HCI  Human-Computer Interaction

HRI  Human-Robotic Interaction

Part I

INTRODUCTION
BACKGROUND

The background chapter is divided into three sections. The first section introduce the reader to our material by giving a short introduction to our subject public displays. The second section describes the problems within the field of public displays that have been handled by this thesis. The last section describes the layout of the entire thesis.

1.1 PUBLIC DISPLAYS

Research around displays that are placed in urban spaces, known as public displays, is a fairly new topic in the field of Human-Computer Interaction (HCI). Today touch and mid-air gestures are the most common modalities for public displays. However, no field studies exist to date regarding public displays combining both. In this thesis we aim to investigate this new aspect of public displays. There is something to be gained from combining both modalities: touch gives precise interaction and tactile feedback, while mid-air gesturing enables interaction from a distance and lets passers-by discover interactivity by interacting inadvertently.

1.2 PROBLEM

The focus of this thesis is to investigate the benefits, drawbacks and consequences of combining touch with mid-air gestures on public displays. We will have a large focus on the behavioral aspects of people interacting with our system. We will also describe the main problems connected to this kind of displays as well as proposing solutions to them. Drawing these conclusions, we want to give guidelines and design recommendations to people who implement and deploy public displays.

1.2.1 Problem statement

How do you combine touch with mid-air gestures on public displays and what are the consequences regarding social behavior and interaction patterns?
1.2.2 Subproblems

In order to solve the main problem of this thesis, we subdivided the problem into several smaller, more concrete ones. In this way we can solve them by designing studies and performing analyses on the different parts.

- How do you communicate touchability to passers-by?
- How do you motivate people to touch?
- How can you attract users to such an application combining both modalities?
- What does the location of the public display matter for how it is used in regard to those two modalities?
- How do these two modalities used influence the honeypot effect and how can we quantify it?
- Does there exist a point where people prefer to stand while performing mid-air gestures?
- Does there exist a similar point where people prefer to stand while using touch modality?

1.2.3 Limitations

In this thesis we will not concern ourselves with investigating possible new interaction techniques combining these two modalities. This has already been done in previous research papers with digital surfaces [6, 3]. We also have some technical limitations narrowing our scope. We will only use single touch technology and due to the limitations of our depth camera, we will not be able to track motions of the user when standing closer than 0.8 meters, since the distance of use lies between 0.8 and 3.5 meters.

1.3 Outline

The thesis is divided into three parts: Introduction, Results and Conclusion and discussion.

Introduction: The introduction part should give the reader basic knowledge about our topic and what we have done. It is composed by four chapters and they will:

- Introduce the thesis and the research field connected to the thesis (see Background and Related Work).
- Describe the problems that have been tackled in the thesis (see Problem).
• Explain how the problems have been tackled in the thesis (see METHODS).

RESULTS: After the introduction part where the research field and the topic of the thesis have been introduced the result parts follows. The result part presents the studies that we have performed on the topic and their results. This part is composed of one theory study and three field studies:

**THEORY STUDY: MODALITIES** discusses the problems that arise while creating a touch and mid-air gestures enabled application.

**FIELD STUDY: LOCATION STUDY** We conducted a series of field studies that compared two affordances for touch, two applications and three locations. These studies resulted in a paper submission to the ACM Conference on Human Factors in Computing Systems 2013 (CHI’13), which is included in its original form. One part of the paper describes the transitions people made, i.e. which modality they started in and how they changed modality during interaction. Another part takes on some social behavior we observed.

**FIELD STUDY: NOTICING TOUCH INTERACTION** describes an investigation of four different conditions of an application. The purpose of the study was to find a correlation between using a silhouette-representation of the user and the passers-by to touch ratio.

**FIELD STUDY: QUANTIFYING THE HONEYPOT EFFECT AND THE N/T POINTS** A follow-up study to the location study mentioned above. The study investigated where users were standing when they made transitions during their interaction. Additionally the study looked into the effect of interactions on subsequent interactions (Honeypot effect).

**CONCLUSION AND DISCUSSION:** This part sums up the results of the thesis and discusses different aspects of the thesis. The discussion chapter also provides a look at future work that would improve and continue the research that was started with this thesis.
This chapter will present the literature review within our field and try to give the reader a good overview of what has been done before.

2.1 COMBINED MID-AIR INTERACTION WITH TOUCH

There exist a number of papers that use combinations of touch with mid-air interaction for different reasons and in different settings.

Marquardt et al. [6] presented some combinations of mid-air hand gestures with touch gestures. They considered the surface and the space above an interactive tabletop as being one continuous interaction space. Some of the suggested interaction techniques included:

- Using a virtual shadow cast from the hand on the table, giving additional feedback.
- Extending the reach of touch gestures by ray casting from the fingertip as a mid-air gesture.
- Continuing a touch-dragging-gesture on the tabletop by lifting up the hand above it and continuing the interaction in mid-air.
- Allowing 6 Degrees of freedom (DOF) manipulation of an object in the air by the movement of the hand, after picking it up from the table.

Hilliges et al. [3] implemented a prototype of a 2D tabletop display augmented by hand gestures performed in the space above. These pinch gestures made it possible to pick up objects and manipulate them in 3D. A virtual shadow rendered under the gesturing hand helped in strengthening the connection between the user’s gestures and the 3D space seen on the tabletop. It also acted as feedback so that the user would see when a pinch gesture was correctly recognized.

Vogel et al. [15] provided an interaction framework from passing-by to personal interaction with public displays. Proxemics (body orientation and position) were used to transition the system between phases from ambient interaction to direct touch manipulation. For personal interaction with their prototype, both mid-air gestures and touch were supported.

2.2 FIELD STUDIES

Field studies observe the behaviour of users interacting with public displays in a natural context.
Peltonen et al. [10] studied the behaviour of people around a large public multi-touch screen deployed in a shop window recording a total of 1199 people interacting. The application was a multi-touch interface used to browse and manipulate images. They found a lot of social patterns such as people negotiating about the space on the screen, conflict management between users, teamwork and turn taking.

In an other field study, Müller et al. [9] focused on how people notice the interactivity of public displays. In this field study they deployed three interactive displays along a shop window. The application was a game designed to cause inadvertent interaction. This worked by the way the image of a passing-by user on the screen interacted with bouncing footballs. They came to the conclusion that it takes some time for a passing by user to notice the interactivity of a screen, the so-called landing effect. They also looked at social behaviour around the screen and found that people often interacted in groups and that it was common for people to form multiple rows in front of the screen.

Ten Koppel et al. [13] studied different configurations of several flat public screens put together (flat, concave and hexagonal). These each produced different actor, audience and passer-by behaviour.

2.3 STAGES OF INTERACTION

A number of research papers describe the different stages of interaction that a passer-by experiences before and after being an active user. In practice, they all describe a sort of funnel, from being a passer-by in the vicinity of the display to an active user. Each step captures some of the users but not all, so that the final stage where some interaction is performed is the ultimate goal to maximize for a designer.

Streitz et al. [12] divided the stages of interaction according to the corresponding 3 interaction zones: Ambient Zone, Notification Zone and Interactive Zone. These are dependent on the distance between the user and the display.

Brignull et al. [1] similarly divided interaction into 3 activity spaces and their activities: Peripheral awareness activities, focal awareness activities and direct interaction activities.

Vogel et al. [15] extend the work of Streitz et al. [12] into a 4 discrete phases interaction framework: Ambient Display, Implicit Interaction, Subtle Interaction and Personal Interaction. They emphasize fluid transitions between the phases and sharing of information between users when not obscuring the screen.

Michelis et al. [8] introduced the audience funnel, in which he instead of the above proposed 6 different phases: passing by, viewing & reacting, subtle interaction, direct interaction, multiple interaction and follow-up action.
2.4 Design Constraints

Triesch et al. [14] formed a set of design constraints within the field of Human-Robotic Interaction (HRI). The following list is a set of those constraints mentioned in his work that we also think are applicable within the HCI field and public displays.

**REAL-TIME**: It is important that the system is able to react on input in real-time, else it will be tiresome to use [14].

**PERSON INDEPENDENCE**: The application should be made Person independent i.e. it should address everybody [14].

**COME AS YOU ARE**: The system shouldn’t require the user to bring anything with them to be able to interact with the system, i.e. come as you are [14].

**NATURALNESS**: The system should provide ways for the user to interact naturally, i.e. there shouldn’t be any learning threshold to be able to interact with the system [14].

2.5 Live Mic Problem

The book *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture* written by D. Wigdor et al. [16] discusses the problem about mid-air gesture suffering from a "live mic" problem. That is: every motion and gesture a user performs in front of a mid-air enabled system is recorded and could be interpreted as an input, even the unintentional ones.

2.6 The Honeypot Effect

The honeypot effect is a social phenomenon that occurs in the vicinity of public displays. When people show interest in a public display, as active users or as audience members, it creates a social buzz around the application [1]. Hereby a social atmosphere is created around the display: people give signals to others that they are open for social interaction. What happens next is a steady increase in usage. Once there is an initial crowd, people will be attracted by it and themselves attract more people [13]. In other words, when someone is using the public display there is a larger chance of people nearby to also become active users. A big crowd in itself is also an attention grabber, people want to know what the buzz is about. This behavior is recorded for keyboard installations [1] as well as for touchscreens [10] and mid-air gesturing [8]. For gestural interaction, it seems that the honeypot both attracts attention and communicates interactivity. When caught in a honeypot effect people sometimes place themselves in multiple rows.
in front of the display, or as an observer try to position themselves so that they are not represented on screen [9].
METHODS

Here we will describe the methods used during our research and discuss their validity and motivation.

3.1 FIELD STUDIES

To find out the answer to our research questions we set out to study people using public displays "in-the-wild" using our own, custom made application.

Field studies are especially valuable in uncovering and validating implicit assumptions about user behavior, which usually are taken for granted in lab environments. There are some main differences between controlled lab studies and in-the-wild field studies. In the former, the participant is brought to a specific place and is given explicit instructions by a researcher. There is always someone at hand to explain the purpose of the application and the task of the experiment. These demand characteristics are almost non-existent in-the-wild, which gives a much more real user experience. Marshall et al. [7] argue that lab studies fail to capture all complexities that occur with users that act in real situations where the technology is ultimately to be placed. Therefore, although being more expensive to perform, field studies have become central to HCI. Although some of the major disadvantages of field studies are: unknown biases in data and no guarantee that the collected data is representative [4], we are convinced that it is the right way to collect this kind of data on user behavior.

3.1.1 Application

To perform our studies in-the-wild, we designed a system called MirrorTouch that combines both touch and mid-air gesture interaction. We used a 65” Samsung touch screen and an Asus Xtion depth camera to support this. Because of factors outside of our control we developed two types of content for MirrorTouch: artistic and commercial (in some locations we were only allowed to deploy commercial content). This gave us an advantage though, since we could now compare different factors of the application design as well. The two applications were both games using physics simulation to motivate people to interact with the system. However, they differed by the content presented (see Figure 1).
12

Methods

(a) MirrorTouch: commercial content  
(b) MirrorTouch: artistic content

Figure 1: Different types of content in the application.

Commercial content: Mid-air gestures were used to play with virtual cubes and earn points. Touching started an advertising screen or allowed the user to skip interstitials.

Artistic content: Mid-air gestures let users interact with virtual flowing water. Touching made a duck appear that floated around.

Both of the contents where also designed so the design constraints mentioned in the related work section 2.4 would be fulfilled. Additionally to those constraints mentioned in Related work, we also formed two own constraints that we think should be taken into consideration when creating a public displays application.

Passing-by interaction: A passer-by should be informed in an early stage that interaction with the system is possible. It’s also important that users have the possibility to come and leave as they want.

Social acceptance: No one wants to be embarrassed in public spaces. The input interaction the system supports should have an social acceptance by the people surrounding it.

Table 1 shows a list over all the design constraints that we took into consideration when creating our application.
3.1 Field Studies

<table>
<thead>
<tr>
<th>Design constraint</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real-time</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Person independence</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Come as you are</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Naturalness</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Passing-by interaction</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Social acceptance</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Design constraints fulfilled

3.1.2 Locations

The following locations were used throughout our different experiments. In each location the public display was placed next to the main walking paths. There were in total four different locations, and next we describe their characteristics, also summarized in Table 2.

Research corridor “the long night of science”: The corridor (Figure 2) was quite narrow and the screen was positioned towards all visitors entering the exhibit. The main audience this night consisted of families with younger children and some couples, seniors and teenagers who were interested in science.

The university corridor: This location (Figure 3) had a quite narrow space which strongly defined the walking path. The corridor lead from an elevator area to a university café on the 20th floor of one of the campus buildings. This café is a popular place for tourists, due to its skyline view over the city. All kinds of people including students come up here to eat breakfast, lunch, have a coffee break or just enjoy the view. The café entrance is the same as the exit door. This means that every person entering came out the same way and passed the screen again. To decrease the number of people using the application for a second time, we turned it so it faced the people entering the restaurant.

The university hall: This location (Figure 4) offered a large space for waiting or walking around. The large hall connects the campus courtyard to the main university restaurant and to different cafés and meeting points. Here also a mix of people spend their time, because it is open to everyone and not only students of the university. For maximum visibility, the screen was placed by the main walking path, although rotated so that people would most often approach it from the side. This was done because the large obstructive screen had to be placed along a wall.

The exhibition hall: This hall (Figure 5) hosted a major consumer electronics trade fair, IFA. It is one of the oldest industrial exhibitions of Germany and today one of the world’s leading trade shows for
consumer electronics and home appliances. Here the display was installed next to a large exit/entrance to the hall, rotated to face most passers-by. The typical visitors of this fair were business people, families, young couples, seniors and a lot of employees of the different companies exhibiting.

### 3.1.3 Logging and observing

For each field study, we collected continuous screen capture and anonymous depth videos of users. The anonymity of the videos was important because of strict privacy laws concerning video capture in Germany. Logs were later automatically searched for situations with users in them and manually annotated. Manual annotation was done to prevent positive errors as the kind where a door is falsely detected as a human by the depth sensor. We also conducted on-site observations (see figure 6) and semi-structured interviews with users, to find out more about their motivations for interacting.
3.2 Statistical Validity

3.2.1 Inter-rater reliability

In all situations where manual annotations of log videos was required, both researchers of this thesis either collaborated on making assessments or split up the log files evenly between them. To make sure there was no bias between the assessments of the researchers, an inter-rater reliability test was performed using Cohen’s Kappa. Results of this test can be seen in Chapter 5. This measure is used to describe agreement between two annotators for categorical data, as in our case. It produces a Kappa value between 0-1 which tells how much they agree, taking into account the agreement occurring by chance [2].

The formula is as follows, where in a $k$ by $k$ confusion matrix $f$, an element $f_{ij}$ defines the number of cases that the first observer assigned to category $i$ and the second to $j$. So, $f_{jj}$ is the number of agreements for category $j$. Then, $P_o$ stands for the observed proportional agreement, $r_i$ and $c_j$ the row and column totals for category $i$ and $j$, and $P_e$ the expected proportion of agreement.
16 methods

Figure 6: Two of the researchers observing from an inconspicuous location.

\[ P_o = \frac{1}{N} \sum_{j=1}^{k} f_{jj}, \quad (1) \]

\[ r_i = \sum_{j=1}^{k} f_{ij}, \forall i, \text{ and } c_j = \sum_{i=1}^{k} f_{ij}, \forall j, \quad (2) \]

\[ P_e = \frac{1}{N^2} \sum_{i=1}^{k} r_i c_i, \quad (3) \]

Finally the Kappa value which measures agreement is obtained from:

\[ \kappa = \frac{P_o - P_e}{1 - P_e}. \quad (4) \]

3.2.2 Chi-square

Many of our studies depended on comparing conditions and noting down effects in form of different actions performed by the user. This gives us nominal (categorical) data instead of the usual continuous numerical data that for example time measurements would give us. The recommended test to perform on this data, for checking the significance, is the Chi-square test [11] for independence. In this case, we want to try the null hypothesis, that is that the occurrence of the observations is equally probable. The Chi-square test gives us a probability of the null hypothesis being true, given by the p value. If the value of p is under 0.05 there is less than 5% chance of the result being due to chance. In the HCI field if the value is smaller, we reject the null hypothesis and claim the effect observed to be significant [5].

One drawback of the test is that it shouldn’t be used in small sample sizes, when the expected frequencies in the experiment are lower than 5 [11]. This was never the case in our studies, so we will use this test throughout the thesis.
Part II

RESULTS
In this chapter, we will describe two things 1) the relevant modalities for input: touch and mid-air gestures and 2) the problems that occur when using them and some possible solutions to them.

4.1 **TOUCH AND MID-AIR GESTURES**

The main difference of touch over mid-air gestures is that touch has a natural delimiter when the user wants to input data (touch) and when not (no touch) [16]. Mid-air gestures in contrast suffer from the *Live mic problem* (see Chapter 2.5). There are three general solutions to this problem:

1. Some special gestures may be reserved as system input. These gestures should then not occur naturally. For example, pinching thumb and index finger may execute a dedicated command.

2. In clutching, a special gesture changes the mode to system input. For example, when the user pinches, the system may count the number of fingers on the other hand to execute a command.

3. Use different modalities to switch between modes. In the “put that there” system, the user could say “put that”, while pointing to an object (speech modality), and then say “there”, while pointing to a new location.

4.2 **STATE-TRANSITION**

One easy way to describe how the different input modalities behave is to use a model based on a state-transition diagram [16]. It is possible

![State-transition model of input](image)
to describe a system with such a diagram if the system is composed of a finite number of states. Figure 7 shows the two different state-transition models for mid-air gesture and touch.

As previously mentioned, touch can naturally distinguish two input states, no touch and touching (see Figure 7b). Users even receive tactile feedback when they switch modes. Naturally, mid-air input does not distinguish states. However, another state can be added through clutching (compare Figure 7a with 8).

4.2.1 Primitives

If only one state is available, a very simple interaction primitive is displacement. Through displacement, one can deform non-rigid virtual objects or apply friction forces to grab things. Either a cursor or a shape may displace virtual objects. This is similar to how we interact in the physical world, which is essentially by displacing physical objects with our fingers/body. If a second state is available, we can also use the interaction primitive of clicking. Like displacement, clicking can be used both to move or manipulate objects. However, instead of displacing an object from two opposite sides, by clicking usually the user representation (e.g. cursor) is placed over a virtual object. When clicked, the object is engaged and is then influenced by the user representation. Clicking enables more accurate (pixel-precise) interactions. Both primitives for user representations are illustrated in Figure 9 for comparison.

4.2.2 Displacement and Clicking

Both touch and mid-air can use displacement and clicking (with clutching) as input primitives. When combining both, four possibilities arise (see Table 3). When either clicking or displacement is used both for touch and mid-air interaction, most likely interaction will be redundant. Thus, the same things can be achieved either by clicking on or displacing the same virtual objects. Clicking mid-air but displacing via touch may make less sense, because touch supports clicking
4.2 State-Transition

Figure 9: Interaction primitives for input. Left: cursor input, right: shape input technique

<table>
<thead>
<tr>
<th></th>
<th>Touch</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Click</td>
</tr>
<tr>
<td>Mid-air</td>
<td>Click</td>
</tr>
<tr>
<td></td>
<td>Displacement</td>
</tr>
</tbody>
</table>

Table 3: Using either clicking or displacement for both modalities mirrors the same functionality for mid-air and touch. The utility of combining touch displacement with distant click is unclear. Combining distant displacement with touch click raises potential problems (see Figure 10).

Figure 10: When the mirror image displaces objects, it is impossible to touch them.

naturally through two states, and mid-air clicking is more cumbersome due to clutching. Combining mid-air displacement with touch clicking may seem the most natural, but creates a problem (see Figure 10). When users displace objects through their mirror image, which is projected on the display with orthographic projection, they are never able to click objects. Whenever they approach an object in order to click it, it is displaced by their mirror image. Even when using a different projection, this problem persists. Objects that the user is just about to click may be displaced by the users themselves or other users.
4.2.3  **Touch and mid-air objects**

One possibility to deal with this situation is to provide two kinds of objects. One kind of objects can be displaced via mid-air gestures, and the other kind of objects can be clicked via touch. One central question is then whether these two classes of objects influence each other (see Table 4). In the case that they do not influence each other, there are basically two different interactions on the same display which are not necessarily integrated. A simple mid-air ball game could for example be overlaid over any existing touch application. The case that mid-air objects influence touch objects may prove problematic for the same reason as above, if a user tries to touch an object which gets influenced by his body movements at the same time. The most elegant solution may be to have the touch objects influencing the mid-air objects, but not vice versa. This way, both kinds of objects can be integrated into the same application without conflicts. Another simple way to avoid the conflict is of course to simply disable mid-air gestures as soon as one user is close to the display. This however may create problems in multi-user situations.

4.2.4  **Our application**

We apply all this theory to our own application design, as described in subsection 3.1.1. Our approach to solve the problems discussed here is simple and straightforward. MirrorTouch with artistic content has only one type of object: mid-air objects. When the touch functionality is used in this content a mid-air object appears (a duck). MirrorTouch with commercial content has a similar solution proposed, where the touch application simply is being overlaid over the mid-air application. In both versions mid-air gesturing functionality is limited to displacement of objects (water/cubes) with the whole body acting as a shape input rather than a cursor input.

<table>
<thead>
<tr>
<th>Mid-air objects influence touch objects</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Touch object influence mid-air objects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>Conflict</td>
<td>Integrated</td>
</tr>
<tr>
<td>No</td>
<td>Conflict</td>
<td>Separate</td>
</tr>
</tbody>
</table>

Table 4: When mid-air objects influence touch objects, the conflict shown in figure 10 can occur. If touch objects influence mid-air objects, but not vice versa, content integrates both modalities. If neither objects can influence another, two entirely separate contents (mid-air and touch) are shown on the same display.
The Live mic problem that D. Wigdor et al. [16] takes up is simply ignored in our application, and all gestures in front of the sensor are interpreted as input to the system.
FIELD STUDY: LOCATION STUDY

The results from a field study aimed at investigating the impact of different locations on user behavior can be read in its original form in Appendix A. Subsequent studies will build upon the results and discussions of this paper.
FIELD STUDY: NOTICING TOUCH INTERACTION

In this chapter we present the studies aimed at understanding what particular variations of our application would make people touch the screen. It turned out that adding mirror based mid-air gesturing capabilities to a touch display increases the usage.

6.1 BACKGROUND & GOAL

A prestudy performed with the artistic content at an exhibition called "The long night of science" (see Chapter 3) revealed that the touch ratio of people passing by was almost zero when no affordance was used for touch. Therefore we decided to further investigate how to make more people touch the screen and discover all modalities. A first step was to find the right affordance for communicating touch, of which the results can be found in Appendix A. In the second step, which we present in the following study, we varied different conditions to investigate how variations of the application impacted passers-by to touch the screen. This result could then be used by for example advertisers or designers of public displays to increase the number of fine grained interactions users make by touching the screen.

6.2 HYPOTHESIS

The hypothesis for this experiment was that the mirror image presented to the user (allowing mid-air gesturing) would be distracting from the fact that the user also could touch the screen. We thought that a display whose only modality was the touchscreen would attract more users to touch.

6.3 SETUP

During two whole days we deployed the artistic content in the university corridor. We were making notes for every passer-by, whether he/she touched or not, and which direction the person was walking. At the same time depth video and logs were captured for later manual and automatic analysis.

The variable we wanted to measure was the user conversion rate, i.e. the percentage of users passing by that stop to make a touch gesture. The independent variables were whether a mirror image of the user was on screen or not, and also the presence of interactive water.
In the design of this experiment, we investigated the effect of mirror interaction on the users ability to notice the screen’s touch capability. In half of the application instances presented to passers-by, the interactive water was removed. This was done to study whether one way of interacting (the water interaction) would discourage the user from searching for more ways to interact with the screen (touching).

In five minutes intervals, cases A to D (see table 5) were randomly cycled throughout the day. Five minutes per condition allowed us to vary and randomize the cases sufficiently throughout the day, for randomization of independent variables. At the same time this allowed multiple conditions to be presented around the same time period of the day, for example during restaurant rush hours.

### 6.4 Results

The results can be seen in Table 6.

Of all conditions, the touch screen with mirror based interaction (i.e. mirror representation interacting with physics) attracted the most passers-by to touch (18.2% of passers-by). This is a substantial improvement over the other conditions, as seen in Table 6.

Although, the second day of studies gave different results, as seen from Table 7.

When merging the two days and looking at them as a whole, we obtained the following results (Table 8):

What’s clear is that adding mirror based interaction and physics gives more interactions overall than touch alone, here shown for both days combined (Table 9).

<table>
<thead>
<tr>
<th>Case</th>
<th>Percentage of passers-by touching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mirror image &amp; interactive water</td>
<td>18.2%</td>
</tr>
<tr>
<td>Mirror image</td>
<td>10.5%</td>
</tr>
<tr>
<td>Water</td>
<td>12.1%</td>
</tr>
<tr>
<td>Touch button</td>
<td>16.7%</td>
</tr>
</tbody>
</table>

Table 6: Day 1: Passers-by touch conversion rate
### 6.5 Discussion and Validity

The semi-structures interviews on twelve users revealed that most explicitly said that they touched the screen out of curiosity (4 out of 7 touching users). We also found that none of the three non-interacting users had observed the screen being there at all. The most common reason for perceiving the touch modality was the label (“touch here”), not the button.

Although some trends in the data seemed convincing, a Chi-squared test revealed that the results of this particular study were not significant ($p > 0.05$).

However, looking at the numbers obtained, most trends suggested our hypothesis to be false. Instead of mid-air gesturing causing distraction from the touch modality, we can see a trend that more ways to interact lead to more touches (Table 8: 15.2%) and to more interactions overall (Table 9: 20.1%). The only condition that broke the trend is the one where nothing but a button was shown on the screen. We

---

#### Table 7: Day 2: Passers-by touch conversion rate

<table>
<thead>
<tr>
<th>Case</th>
<th>Percentage of passers-by touching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mirror image &amp; interactive water</td>
<td>13.1%</td>
</tr>
<tr>
<td>Mirror image</td>
<td>12.0%</td>
</tr>
<tr>
<td>Water</td>
<td>15.7%</td>
</tr>
<tr>
<td>Touch button</td>
<td>18.0%</td>
</tr>
</tbody>
</table>

#### Table 8: Day 1 & 2 combined, passers-by touch conversion rate

<table>
<thead>
<tr>
<th>Case</th>
<th>Percentage of passers-by touching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mirror image &amp; interactive water</td>
<td>15.2%</td>
</tr>
<tr>
<td>Mirror image</td>
<td>11.3%</td>
</tr>
<tr>
<td>Water</td>
<td>14.0%</td>
</tr>
<tr>
<td>Touch button</td>
<td>17.4%</td>
</tr>
</tbody>
</table>

#### Table 9: Day 1 & 2 combined, passers-by interaction conversion rate

<table>
<thead>
<tr>
<th>Case</th>
<th>Percentage of passers-by interacting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mirror image &amp; interactive water</td>
<td>20.1%</td>
</tr>
<tr>
<td>Mirror image</td>
<td>16.4%</td>
</tr>
<tr>
<td>Water</td>
<td>14.0%</td>
</tr>
<tr>
<td>Touch button</td>
<td>17.4%</td>
</tr>
</tbody>
</table>
attribute this rather large number of interactions to the curiosity effect. When the use of such a large display was as unclear as in this situation, people were puzzled about its purpose and interact out of pure curiosity. We could observe a lot of people standing in front of the empty display, scratching their heads as a duck dropped to the bottom. This effect was supported by our interviews, where curiosity was the most common motivation for touching the screen.

In conclusion we want to encourage owners of large touch displays to add mid-air interaction possibilities as well as it attracts more attention from passers-by and increases the number of interactions overall. We also suggest to add an element of curiosity to capture the interest of passers-by.
In this chapter we present the results of our studies focused on where users of touch and mid-air gesture enabled public displays position themselves, and also on the effect of interactions on subsequent interactions (honeypot effect). The result was that while people position themselves at arm’s length for touch interaction, they stand back to position themselves in the middle of the screen for mid-air gesturing. We also present a successful measuring of the honeypot effect.

7.1 Background & Goal

While Chapter 5 answers the question of how people make transitions between touch and mid-air gesturing, this study aims at also describing where these transitions take place. In the works of Ten Koppel et al. [13] the aspects Nimbus and Focus are discussed. These are related to the area where a user can perceive content on the display and the area where interaction is possible. For public displays which are both touch and mid-air gesture enabled we would like to extend this description to include the preferred point of interacting for both modalities.

The second part of this study investigates the honeypot effect. See Chapter 2.6 for a detailed description of the effect and its definitions. It has been discussed in a lot of related work such as [1, 13, 8, 9], but there has never been a study which quantifies the effect. This is why we aimed at understanding the differences between touch and mid-air gestures when it comes to the honeypot effect, and also put a real number on the actual strength of the observed effect.

7.2 Hypothesis

Our assumption was that the honeypot effect would produce a significant result in our log files. We also thought that there would be distinct point where people stand when they make the transitions explained in the field study of Chapter 5.

7.3 Setup

The setup was the same as our other studies, and was performed in the university corridor. We analyzed the log files by taking the center of gravity for each user and mapping it to the 2D floor plane. Then
we divided the space in front of the screen into discrete bins. Each frame where a user was in front of the screen was represented by one data point in one of the bins. In this way we created a heatmap of where users were spending the most time (frames) in front of the display.

7.4 RESULTS

We now present the results in two parts: where people interacted (positioning) and when people interacted (honeypot).

7.4.1 Positioning

The results can be seen from the following heatmaps in Figures 11, 12, and 13. The range varies from dark blue color to red for most activity in front of the display. The first image (Figure 11) can be seen as a baseline for this particular area (University corridor) since it shows all activity during 3 days of logging user activity. The warmer colors showed up around the edges and the main walking paths. Only users which spent less than 5 seconds in front of the display were considered, who didn’t have touch events registered in this time period. In Figure 12 we instead looked at users touching. The data was filtered by only looking at users with touch events registered and who spent more than 5 seconds in front of the display. In the third image (Figure 13) we gathered data from gesturing users, filtered by looking at people who spent more than 5 seconds in front of the display, without touch events.

Figure 11: Heatmap of all passers-by, university corridor (as seen from above)
### 7.4.2 Quantification of honeypot effect

In order to quantify the honeypot effect we analyzed the log files by looking at interactions that took place up to eight seconds from a previous one. Eight seconds was chosen to be a reasonable period of time while keeping the number of false positives down. To measure the effect, we first established a baseline where we took the average conversion rate for a passer-by to interact with the display. Then we compared this side-by-side with the conversion rate obtained from these honeypot timeframes around eight seconds after an interaction occurred. In Diagram 14 the honeypot effect is split up into different types of modalities that preceded the current interaction. A Chi-squared test revealed ($p < 0.05$) that there was a significant effect of the used modality on the chance of the subsequent interaction (52.2%)}
34. FIELD STUDY: QUANTIFYING THE HONEYPOT EFFECT AND THE N/T POINTS

Figure 14: Conversion diagram with different preconditions. The first plot is the baseline where all passers-by during three days are measured. The last three measure interactions within the predefined honeypot timeframe for both modalities, for touch alone and for mid-air gesturing alone. Based on 1701 passers-by.

Figure 15: Histogram over the group sizes that were seen interacting during a two day observation of the university corridor. Based on 81 interaction sessions.

gesture conversion rate vs 28.9% for touch compared to the baseline of 17.3%).

7.5 DISCUSSION AND VALIDITY

As we suspected, two clear interaction points could be distinguished. These points were also confirmed by our observations, and we will call them the T (touch) and N (mid-air gesturing) points. People who used both modalities made clear transitions between these two discrete points. The point for mid-air gestures seems to be a little off from the center of the screen. This could be explained by the fact that
people often don’t want to be in the way of others, and step aside from the main walking path while interacting.

Also, the honeypot effect was even more apparent than we hoped for. There is a substantial increase in how many people interacted following up on a previous interaction. However, this could be explained by the fact that what we were measuring in this case also includes the "group effect", i.e. people don’t always interact alone. According to other results in our studies (see Figure 15) the average group size of people interacting with the screen was 1.9 at the University Corridor. Interesting to note is also the fact that mid-air gestures clearly and significantly increased the honeypot effect for public displays, compared to touch gestures. This is again a clear example of the strength of mid-air gesturing: drawing attention and promoting social and expressive behavior.
Part III

CONCLUSION AND DISCUSSION
CONCLUSION

We conclude by summarizing the design guidelines and results from previous studies in the following points:

- Different modalities of public displays are mainly communicated to users by the use of good affordances. In the case of mid-air gesturing this means enabling inadvertent interaction by on screen representation. For the touch modality this means an affordance in form of an explicit call-to-action text on screen. Designers of public displays should use both for maximum conversion rate and usage of both modalities. (See Chapter 5)

- A main factor in motivating people to interact is curiosity. This should be taken into account when designing content for public displays. (See Chapter 6)

- A main factor in attracting people to interact are good affordances for using the modalities of the display and the use of interesting content that is responsive. For example a large number of physical objects that are interactive on screen are assumed to increase the conversion rate along with clear affordances. (See Chapter 5)

- The location where the display is placed has a significant effect on what order the modalities are used in. This should be taken into account when designing the interaction. (See Chapter 5)

- We observed a significant increase of usage when the honeypot effect occurs (106% increase) and the honeypot effect for mid-air gesturing is significantly larger (52.17% conversion rate) than for touch gestures (28.89% conversion rate). (See Chapter 7)

- There seems to be a certain point where people preferably stand to make mid-air gestures and where they stand to touch public displays. For the former this is in the middle of the screen, away from the main walking path and for touch this is at arm’s length. (See Chapter 7)
DISCUSSION

In this chapter we look back at the work we have done, and ponder on what possible new directions our thesis could take.

9.1 REFLECTION

This thesis about a completely new topic in an unknown environment was a challenge. Due to the very friendly atmosphere at Quality and Usability Lab of Deutsche Telekom Laboratories, we could adjust quickly to the task we were supposed to do.

The conclusions we have made in this thesis have been aimed to be the most general and logical derivation from the results we got from our field studies. However, we are aware of one drawback with the results from our field studies, and that is that mostly students tested our application. This has unfortunately been out of our control since it requires permission to put up a public screen for field studies.

9.2 FUTURE WORK

While working on our thesis a number of new project within the field of HCI and public displays came to our attention:

CONTINUOUS INTERACTION WITH THE BODY: Asus Xtion depth camera has a limitation when it comes to how close a body can be tracked. One interesting project to work on would be to combine a set of depth cameras to avoid this problem and track the body all the way to the display surface. This technology would make the interaction more interesting since a touch point would converge with the mirror-representation of the human body.

ONE SCREEN EACH FOR THE MODALITIES: We combined the two modalities on one screen during our project, but the two modalities also make it possible to have two separate screens: one for mid-air interactions and one for touch interactions.

DEEPER INTERACTIONS: Our application was simple and we kept the interaction as simple as possible for both modalities. Instead, for mid-air gestures, proxemics could be used for deeper and more advanced interaction. For the touch interaction more advanced interactions than just creating mid-air objects or changing between interstitials could be added.
Part IV

APPENDIX
MIRRORTOUCH LOCATIONS STUDY: A RESEARCH PAPER SUBMITTED TO CHI’13
MirrorTouch: The Impact of Location on Audience Behavior for Touch and Mid-air Gestures on Public Displays

1st Author Name
Affiliation
Address
e-mail address

2nd Author Name
Affiliation
Address
e-mail address

3rd Author Name
Affiliation
Address
e-mail address

Figure 1. a) A user performs a transition from mid-air gestures to touch and back to mid-air gestures again. b), c) and d) Conversion diagrams for transitions in two locations (corridor vs. hall) and for two contents (artistic vs. commercial), each one based on around 350 passers-by.

ABSTRACT
We present results from four field studies aimed at investigating the impact of location on user behavior for touch and mid-air gesture enabled public displays. These field studies, were conducted using MirrorTouch, a public display combining touch and mid-air gestures. The field studies compare two affordances for touch, two applications and three locations. Our contributions are twofold. First, we show how affordances, application and location impact the conversion rate. A call-to-action causes more users to touch than a button. Small variations in the application lead to a user increase of 390%. Location impacted which modality was used first. Second, we qualitatively describe users’ behavior and derive guidelines for designers: 1) Design for subtle waiting areas where users do not wait in a queue. 2) The degree of social control should be considered in order to predict expressive behavior. 3) Design both for users and for employees.

Author Keywords
Public displays; field studies; touch; mid-air gestures

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

General Terms
Human Factors; Design; Measurement;

INTRODUCTION
Recently, a number of field studies have explored audience behavior of touch or mid-air interaction with public displays [2, 4, 5]. However, the combination of touch and mid-air gestures has not yet been explored in the context of public displays.

We propose MirrorTouch, a public display combining touch and mid-air gestures. We used a 65” touch screen and an Asus Xtion depth camera to support these two modalities. MirrorTouch allows precise interaction and tactile feedback (touch capability). At the same time MirrorTouch lets users interact from a distance and lets passers-by discover interactivity by interacting inadvertently [4] (mid-air gesture capability).

As in other works [1, 7], MirrorTouch combines touch and mid-air gestures. However our work focuses on first time users without prior knowledge (rather than expert users [1]) and public spaces (rather than an augmented reality room [7]).

We deployed MirrorTouch during four field studies to investigate the impact of three locations and two different applications on user behavior for touch and mid-air gestures on public displays. Our main findings are:

• Clear affordances for touch are necessary when mid-air gestures are present.
• Locations impact the type of modality used first.
• Small variations in the application can have a large impact on the number of interactions.
• Subtle waiting situations influence audience behavior.
• Socially appropriate behavior varies strongly with location.
• Permanent residents (employees) influence audience behavior.

From our findings we have derived guidelines for designers, practitioners and researchers of public displays that want to deploy public displays combining touch and mid-air gestures.

Submitted to CHI’13.
Do not cite, do not circulate.
APPLICATIONS AND LOCATIONS

In order to investigate touch and mid-air gestures on public displays, we implemented two variations of MirrorTouch and deployed them in three different locations.

**Applications:** The two applications were both games using physics simulation to motivate people to interact with the system. However, they differed by the content presented. The first application had commercial content (Figure 3a). Mid-air gestures were used to play with virtual cubes and earn points. Touching started an advertising screen or allowed the user to skip interstitials. The second application had artistic content (Figure 3b). Mid-air gestures let users interact with virtual flowing water. Touching made a duck appear that floated around.

**Locations:** We deployed the two systems described above in three different locations. In each location the public display was placed next to the main walking paths. The university corridor (Figure 2a) had a quite narrow space which strongly defined the walking path. In contrast, the university hall (Figures 2b, 4b) offered a large space for waiting or walking around. Finally, the exhibition hall (Figures 2c, 4a) hosted a major electronics trade fair. Here the display was installed next to a large exit/entrance to the hall.

While the audience consisted mainly of students for the university corridor and hall, there were a lot of business people at the exhibition hall.

**METHODS**

We conducted four field studies. Study 1 aimed at defining an efficient affordance for touch interaction that was to be used for the subsequent studies. Study 2 and 3 investigated the impact of location and application on user behavior for touch and gestural interaction on public displays. Finally, study 4 qualitatively investigated the social behavior of people around MirrorTouch.

For each field study, we collected continuous screen capture and anonymous depth videos of users. Logs were automatically searched for situations with users and manually annotated. We also conducted on-site observations and semi-structured interviews with users. Interrater reliability for the two researchers observing interaction sequences was “almost perfect” (Cohen’s Kappa = 0.82, 95% CI = (0.689, 0.954)).

**FIELD STUDY 1: AFFORDANCES FOR TOUCH**

While showing a mirror image of the user has been shown to be an efficient affordance for communicating mid-air gestures on public displays [4], we are not aware of any comparisons between touch affordances. Therefore we conducted a one day field study with 66 users and 697 passers-by to compare two touch affordances (see Figures 3c, 3d). While the first was a traditional button similar to those commonly used on other touch surfaces, the second one was a more explicit call-to-action. The study was performed in the university corridor with the artistic content. The two conditions were switched every 5 minutes.

**Results:** A Chi-square test revealed that the percentage of passers-by utilizing the touch screen significantly differed with the affordance used ($\chi^2_{1, N=697} = 20.9, p < .05$). Call-to-action affordance had a 15% touch conversion rate, while the button had 5%.

**Design Recommendations:** The call-to-action not only had a better conversion rate but it also required less screen real estate. Therefore we encourage practitioners to use a similar graphic to communicate touch technology.

We used this touch affordance (call-to-action) combined with the users’ mirror image for mid-air gestures in the subsequent field studies of this paper.

**FIELD STUDY 2: IMPACT OF THE APPLICATION**

We conducted a second field study comparing two different contents and keeping the location (university hall) constant. The contents were deployed during one day each, giving a total of 53 users and 700 passers-by.

**Results:** A Chi-square test revealed that the percentage of passers-by utilizing the public display significantly differed with content ($\chi^2_{1, N=700} = 20.5, p < .05$). The artistic content increased the number of users by +390%. Furthermore,
the content did not have a significant effect on which modality people chose to use first, nor did it effect which one they used last. The conversion diagrams are shown in Figures 1c and 1d.

We also observed that some users who used touch first never noticed their mirror image (the feedback for mid-air gestures). As the appearance was not in the users’ field of view - in close proximity to the touch location - users did not notice the gestural modality. Similarly, they failed to perceive some effects of their touch interaction when the effects were not shown very close to the touch.

Design recommendations: We learned two things from this study. First, a variation of content can lead to a huge difference in the conversion rate. This is surprising as the two applications were quite similar: both are 1) a game with 2) physics simulation, 3) they also had the same affordances and 4) the same modalities (touch and mid-air gestures). There are two likely explanations for this behavior. One: the artistic content had a larger number of physical objects (particles), which continuously responded to the gestures of the user. This may have lead to a greater feeling of responsiveness [6], and attracted more users. Two: the artistic content looked less polished than the commercial content and may have aroused curiosity as it was not apparently useful for anything. It may also be that passers-by immediately classified the commercial content as “advertising” and ignored it.

Second, designers should include located feedback for both mid-air gestures and touch. Touch requires users to interact in proximity of the display. Thus when the screen is large (65” in our case), users can not have the entire screen in their field of view and can miss the effect of mid-air gestures or their touch if it is not located nearby.

We reused the artistic content for the subsequent field studies.

FIELD STUDY 3: IMPACT OF THE LOCATION

We conducted a third field study to investigate the impact of location on user behavior for touch and mid-air gestures. More precisely, we investigated the impact of location on the transitions between the two modalities. The study was performed with the artistic content in the corridor and university hall during two days with 117 users and 718 passers-by.

Results: A Chi-square test revealed that the first used modality significantly differs with location ($\chi^2_{1, N=117} = 3.9, p < .05$). Passers-by mainly used mid-air gestures as a first modality in the hall (7% vs. 5%) while they mainly used touch first in the corridor (12% vs. 8%). The two conversion diagrams are illustrated in Figures 1b and 1c.

Design recommendations: We implicitly assumed that users would perform mid-air gestures first, regardless of the location. However, we observed that the distance from the screen to the main walking path strongly influenced the choice of the first modality. People passed by the display quite close in the corridor setting (narrow space). Because of this they could easily touch the display. In the hall setting on the other hand, people often made a larger detour to touch the screen (spacious location). Designers can thus influence the proportion of touch/mid-air gesturing by varying the gap between the display and the main walking path.

FIELD STUDY 4: EXHIBITION HALL

We had the opportunity to deploy MirrorTouch at a very large trade fair, which saw a total of 240 000 attendants during six days. This location was radically different from the two previous locations because of the type of audience (business people/customers/employees vs. students/university staff) and because of the context (a lot of advanced technology was in the vicinity). This opportunity provides a complementary location to investigate touch and mid-air gestures on public displays. Despite these advantages, we were constrained to only show the commercial content in this environment.

In this section we qualitatively compare this location to the others previously mentioned, regarding touch and mid-air gesturing. More specifically, we highlight 3 behaviors: 1) waiting behavior, 2) socially appropriate behavior and 3) resident influenced behavior.

Waiting Behavior

In all three locations we observed queuing situations. In the university corridor, the queue of the neighboring small cafeteria sometimes extended all the way to the screen. At the exhibition hall, there was a queue at the counter nearby the screen. Close to the screen in the university hall there were queues at ATMs. Even though just waiting, none of those people did interact with the display. The waiting situations that actually influenced usage were much more subtle. For example, people started to play while waiting for their friends standing in the ATM queue. Similarly, in the university corridor people were waiting for the nearby elevators. Many people walked to the elevators, saw that they had not arrived yet, and walked back to play with the screen. Certainly, as soon as the elevator ringed to announce its arrival, people hurried to get into the elevator (although a few missed it).

Design recommendations: Today many non-interactive displays are installed close to waiting queues [3]. Because of the striking differences regarding the number of interactions, we recommend to instead place interactive public displays close to waiting areas where people are not forced to wait at a specific location.
Socially Appropriate Behavior

We observed that socially appropriate behavior differs between the locations, especially both university areas compared to the exhibition hall. While players behaved “well” and according to what they thought the display was designed for in the exhibition hall, they were engaged in much more playful behavior at the university. For example, two users joined their silhouettes and mimed a camel passing by (Figure 5b).

The difference in behavior was also distinct for children. While children generally behaved “well” at the exhibition hall, a group of children displayed rather extreme behavior at the university corridor. Because there were no people in the corridor, the children seemed to feel unobserved. At one point, they started pushing the screen so hard when “touching” it that it started to shake notably. From our observations, after mid-air gesturing expressively, the touching on the screen also became more aggressive (Figure 5c).

Design recommendations: The location can modify user behavior. More precisely, the degree of social control (for instance, the presence of employees at the trade fair) can elicit “appropriate” behavior. In conclusion, the degree of social control for each location should be taken into account by designers, to predict expressive behavior and to build systems that are physically robust enough.

Resident Influenced Behaviour

We observed that the behavior of passers-by and users changed depending on the people around. For instance an employee was in charge to present the screen during the two first days of the trade fair. This decision was unfortunately not under our control. However, it revealed interesting user behavior for both the employee and passers-by in comparison to when he was not there. This employee played with the screen and turned around when some people passed by to establish eye contact. He then invited them to interact. We observed that he convinced many people to approach and interact with the display. This was also confirmed by our interviews on site.

Additionally we observed that users tended to mimic the same mid-air gestures that the employee performed before (Figure 6). As another example, there were also a considerable number of employees performing surveys with passers-by (Figure 4c). When they positioned themselves next to the display, the number of interactions reduced considerably. Finally, employees of nearby stands regularly played with the screen when they had no customers, trying to beat the high-score (Figure 5a). At the other locations we had a lot of curious employees who sometimes played, but they did not affect others’ behavior noticeably.

Design recommendations: From our observations, people (employees) around the displays can strongly influence users’ behavior. Some employees increase the conversion rate explicitly by addressing passers-by or implicitly by playing with the system themselves. Other employees can reduce the conversion rate by effectively blocking the screen. Therefore designers should design systems not only for passers-by but also for employees, for instance by introducing a high-score.

CONCLUSION

From the insights gained during our field studies, we conclude by summarizing two derived design guidelines: 1) Use strong affordances both for touch (e.g. call-to-action) and mid-air gestures (e.g. inadvertent interaction). 2) Test a prototype of the application in the final location. Experiences from one location or application may not necessarily generalize towards others. In particular, pay close attention to a) subtle waiting situations, b) social control at the location, c) the influence of permanent residents.

REFERENCES


[15] Daniel Vogel and Ravin Balakrishnan. Interactive public ambient displays: transitioning from implicit to explicit, public to

På svenska

Detta dokument hålls tillgängligt på Internet – eller dess framtida ersättare – under en längre tid från publiceringsdatum under förutsättning att inga extra-ordinära omständigheter uppstår.

Tillgång till dokumentet innebär tillstånd för var och en att läsa, ladda ner, skriva ut enstaka kopior för enskilt bruk och att använda det oförändrat för ickekommersiell forskning och för undervisning. Överföring av upphovsrätten vid en senare tidpunkt kan inte upphäva detta tillstånd. All annan användning av dokumentet kräver upphovsmannens medgivande. För att garantera äktheten, säkerheten och tillgängligheten finns det lösningar av teknisk och administrativ art.

Upphovsmannens ideella rätt innefattar rätt att bli nämnd som upphovsman i den omfattning som god sed kräver vid användning av dokumentet på ovan beskrivna sätt samt skydd mot att dokumentet ändras eller presenteras i sådan form eller i sådant sammanhang som är kränkande för upphovsmannens litterära eller konstnärliga anseende eller egenart.

För ytterligare information om Linköping University Electronic Press se förlagets hemsida http://www.ep.liu.se/

In English

The publishers will keep this document online on the Internet - or its possible replacement - for a considerable time from the date of publication barring exceptional circumstances.

The online availability of the document implies a permanent permission for anyone to read, to download, to print out single copies for your own use and to use it unchanged for any non-commercial research and educational purpose. Subsequent transfers of copyright cannot revoke this permission. All other uses of the document are conditional on the consent of the copyright owner. The publisher has taken technical and administrative measures to assure authenticity, security and accessibility.

According to intellectual property law the author has the right to be mentioned when his/her work is accessed as described above and to be protected against infringement.

For additional information about the Linköping University Electronic Press and its procedures for publication and for assurance of document integrity, please refer to its WWW home page: http://www.ep.liu.se/

© Niklas Hillgren, Thor Bossuyt