

Fabian Hennecke, Sebastian Boring, Bettina Conradi, Sara Streng,
Raphael Wimmer, Alexander Wiethoff, Andreas Butz (Editors)

Interactive Surfaces

Hauptseminar Medieninformatik SS 2009

Technical Report
LMU-MI-2009-1, Jul. 2009
ISSN 1862-5207



University of Munich
Department of Computer Science
Media Informatics Group

Fabian Hennecke, Sebastian Boring, Bettina Conradi, Sara Streng,
Raphael Wimmer, Alexander Wiethoff, Andreas Butz (Editors)

Interactive Surfaces

An overview of current trends, developments, and research
in Interactive Surfaces

Preface

This report provides an overview of current applications and research trends in the field of interactive surfaces in terms of hardware (e.g. tables) and interaction techniques.

During the summer term 2009, students from the Computer Science Department at the Ludwig-Maximilians-University in Munich did research on specific topics related to interactive surfaces and analyzed various publications. This report comprises a selection of papers that resulted from the seminar.

Each chapter presents a survey of current trends, developments, and research with regard to a specific topic and includes theories, methods, and findings from hardware design, interaction design and many more. Therefore, the report is targeted at anyone who is interested in the various facets of interactive surfaces and related interaction techniques.

In addition to this report, there are slides from the students' talks available at

<http://www.medien.ifi.lmu.de/lehre/ss09/hs/> .

Munich, July 2009

The Editors

Fabian Hennecke, Sebastian Boring, Bettina Conradi, Sara Streng, Raphael Wimmer,

Alexander Wiethoff, Andreas Butz

Contents

Franz Berwein	
Superimposed Displays	1
Sascha Gebhardt	
Superimposed Displays	9
Miriam Kranz	
Applications for interactive public displays	16
Dhana Sauernheimer	
Personal and public workspaces in interactive surface environments	23
Stefan Grabs	
Hybrid Interaction on Interactive Surfaces	30
Renata Willi	
Interaction Metaphors on Interactive Surfaces	38
Felix Lauber	
Interactive Displays for Creativity Support	46
Sebastian Löhmann	
Applications for interactive public displays	54

Superimposed Displays

Franz Berwein

Abstract— The term "superimposed displays" has been coined to refer to any configuration of two displays where one is superimposed on the other. This applies to setups like handhelds and large context displays, but also head-mounted displays that rely on reality as superimposed display. Recent research concerning this topic has produced a vast amount of papers containing interesting findings that could prove useful for related projects. Yet a comprehensive and focused overview is missing. This paper shall provide a survey of the origin, technologies and appliance of superimposed displays. It will point out different aspects like connection, tracking, projection, navigation and other forms of interaction such as selection or drag-and-drop tasks. Ways to overcome or at least reduce technological obstacles will be highlighted in the course of that. Then the paper will divide the issue at hand in three categories – touch-based interaction, options to extend the interaction space beyond the surface and applications in the domain of augmented reality. For each category the most salient and influential works will be introduced to portrait the state of the art and show the benefits and difficulties of superimposed displays.

Index Terms— Superimposed displays, interactive surfaces, spatially aware, tabletop, handheld, magic lens, peephole displays, mobile, ubiquitous computing

1 INTRODUCTION

Techniques to enrich our visual experience and aid our eyes have been a goal of mankind for a very long time. The dream of interactive environments has been the topic of many movies and an impetus for many scientists. This vision has started to become reality. While the hype over VR helmets at the beginning of the 90s has long vanished, sophisticated new technologies have since gained drive, like interactive tables and walls and tangible surfaces. Today we are at the brink of a merge between reality and virtual reality. Furthermore, we have started to explore the benefits of augmenting reality with computer vision. The last decade saw a fast-progressing evolution in this matter. The topic first gained broader attention in 1993 with the ground-breaking works about ubiquitous computing by Weiser [71] and situated information spaces by Fitzmaurice [17]. While the former inspired innumerable ambitious projects and dubbed the dawning computer era, the latter paved the way for future research concerning interactive surfaces.

Superimposed displays are a sub-genre of interactive surfaces and spatially aware displays. They try to enrich the interactive experience by using well-known metaphors like toolglasses and magic lenses [10], peepholes [80] and annotations. The term "superimposed displays" has been coined for environments where one display is beneath or behind another and the closer display provides an enhanced view of the superimposed area of the remote display. Weiser referred to this setup as *tabs* and *boards* [2, 71]. Often this has been narrowed down to spatially aware handhelds and large situated displays. While this is certainly the predominant field of use, it can not be deemed sufficient to include all applications utilizing this setup. Reminiscent of Weiser's original vision, the term thus tries to broaden the topic to all environments featuring two interconnected displays.

2 MOTIVATION

The motivation for superimposed displays is manifold. Concerning tabletops, the usual way to interact directly means pressing the finger against the surface. This input serves well for location and low-level manipulation (for example, dragging items), but meets its limits when it comes to more sophisticated tasks, especially manipulations of small areas. This enforces the wish for a more intelligent input device that matches these needs and offers fine-grained control. Earlier concepts

involve the metaphor of *focus + context* [6]. The actual system featured a large low-resolution display and a smaller high-resolution display fixed in the center of the large context screen. Users have to pan the entire image to focus a specific area. The disadvantage of this approach is the inevitably stationary setup. Handhelds were then used to combine a movable focus area with a larger context display in the background. This is helpful in two ways: While handhelds naturally provide a relatively high resolution on a smaller display and thus are perfectly suitable for this task, the large display can show its content without the interference of a stationary focus screen. This setup may still be enhanced for some purposes by making the personal display spatially aware. That means location data is automatically transmitted and updated on the personal screen enabling the user to interact (for example, scroll) with the large display by moving his personal handheld device.

Since ubiquitous computing more and more turns from vision to reality and since cellphones have spread all across the globe, small handhelds have become an everyday item worn by everyone. This development offers lots of opportunities, as the notion of superimposed displays could thus become truly ubiquitous. Everybody could use his cellphone to interact with large displays easing his everyday life. Thereby interaction between humans could be enriched greatly – and all of that just with more or less basic, widespread and cheap hardware on the client side if adequate technologies are developed and applied. Developers have already reached far in this matter, but several issues still bind the evolution. To name two major ones: there has yet to be found a solid way to connect large display and handheld and robust means to track the handheld(s) as well, especially when it comes to locating them above the surface. Various solutions have been suggested for both issues and shall be detailed along with other issues that arise in the production of superimposed display solutions.

3 THEORY

Over the past years there has been a growing effort to develop superimposed displays, be it for augmenting reality, better visualization of certain spots on large displays or interaction with objects on them. There is a lot of potential, but also partly unresolved pitfalls that need attention. In the following sections, the major technical issues shall be detailed. These are the connection of the devices, subsequently tracking them, which projection mechanisms to use, navigational issues and other forms of interaction such as selection and text input.

3.1 Connection

To start interaction between two displays, they first have to be connected. This goal can be reached in a variety of ways. There are projects that establish a connection detecting synchronous events like

-
- Franz Berwein is studying Media Informatics at the University of Munich, Germany, E-mail: berwein@cip.ifi.lmu.de
 - This research paper was written for the Media Informatics Advanced Seminar on Interactive Surfaces, 2009

bumping devices together [24, 30] or pressing the same key simultaneously [55]. Another option is to display a random key on the handheld that has to be typed in to synchronize with the large display. This key may be an alphanumeric string, a sequence of motions or gestures or a visual pattern [3, 5, 50, 63].

Connectivity and device identification can also be provided by radio frequency identification (RFID) or near field communication (NFC) tags and a corresponding reader [15, 51, 60]. Although these tags are fairly inexpensive and provide accurate identification, there is a major drawback, namely the need to equip the mobile handhelds with these tags. This detour prevents the envisioned ad-hoc approach in everyday life. Another problem is the reading range applied: Either a small compulsory range is given, limiting the user to a very narrow physical space. Alternatively, a long-range reader may be installed, prompting the user with the difficulty "to judge whether a given tag is within reading range" [76]. This disambiguity may possibly lead to unintended connections and security and privacy concerns.

Bluetooth can also be used to connect two or more devices. One should bear in mind though that only up to eight users in one piconet are supported. It is frequently applied, for example by *LightSense* [46], to transmit data from and to the handheld. It is remarkable though that it takes a considerable amount of time to achieve a Bluetooth connection. Thus it should not be used in time-critical environments. *BlueTable* relies on Bluetooth for this purpose, too, but proposed a different mechanism to initiate a connection [76]. It employs vision-based handshaking to assert proper identification. First the placement of the handheld device on the tabletop is detected by visual means. Then a Bluetooth connection is attempted to all devices in a given range. If they run a specific software (reported by the Bluetooth service GUID), each of them is subsequently commanded to blink its infrared port. If the currently blinking device does not match the position located before, the next device is commanded to blink until the right one is reached. After this initial handshake, communication via Bluetooth may commence. The infrared blinking could also be substituted by a recognizable modification of the handheld screen. The connection and identification can be secured and disambigued by blinking in a unique pattern. The drawback of these methods is that using infrared light is prone to occlusion and therefore the system might erroneously think that a device has been removed while it is only occluded by a hand or an item on the table surface. This can be overcome by building in tolerances for long dropouts.

It is also imaginable to employ other wireless techniques such as WLAN (IEEE 802.11). It features a higher data rate and entails less constraints, but power consumption and hardware prices are higher. Despite of that, *Ubiquitous Graphics* [61] used an IP-based adhoc wireless network where each device acted as an independent peer and synchronized with others in proximity by broadcasting ping signals. This approach may suit well in an environment with tablet PCs, but may not be applicable to smaller handheld devices that are shipped with less capable batteries and storage space.

3.2 Tracking

Tracking the mobile devices connected is the second challenge that research still has to face. RFID or NFC tags have already been mentioned. Another widely used concept is tracking the infrared (IRDA) port of the handheld (active). Passive infrared tracking can be realized with retro-reflective spots on the handheld. This reduces the already low power consumption to zero. The obvious drawback of this technique is the need for an installed infrared device which has become less common on cellphones over the last few years. Another disadvantage is the demand for a free line of sight between the ports and the considerably small range (max. 1 meter) [76]. Bluetooth, developed as an alternative to infrared, may be used to track devices, but is not suitable for the needs of superimposed displays since the resolution of location detection only scales down to room-level [21, 14]. It might be (and is frequently) used for the exchange of data that are not critical in terms of time.

With the rise of built-in digital cameras, the need for flashes has grown accordingly, fueling the inclusion of some kind of flashlight

in modern cellphones. (Bright) LCDs are also a common component of cellphones nowadays. These light sources can be used as another means to track the devices, as *LightSense* [46] has shown. A similar approach has been undertaken by Miyaoku et al. with *C-Blink* which employs hue differences (due to radical switches) on the mobile screen as positioning signal [38]. As emitting light (and using Bluetooth) consumes more power and because power naturally is scarce and precious on mobile devices, this option may only be useful in specific circumstances (for example, short interaction).

Yee [80] mentions several other options for tracking such as optical tracking which means affixing the innards of an optical mouse to the handheld to track its motion on a flat surface. He also suggested inertial tracking, that means tracking of the acceleration and other physical forces that are applied to the handheld. The iPhone provides a glimpse onto the possibilities of this technology which has yet to be tested concerning tracking. Furthermore, he tested ultrasonic tracking, but concludes that its position reading is too slow and tracking is too instable (or interaction too cumbersome) as the transmitter has to be exactly co-planar with the receivers. To overcome these boundaries he hence used fishing lines that ran through two reference points and then were connected to an optical mouse on a flat surface and, on the other end, connected to the handheld. Tension was maintained by two small weights at the end of the tethers. This prototypical construction made it possible to navigate relatively freely while being tracked in a robust fashion. Light, IR and tagged tracking have helped to ease these constraints, but yet a robust and cheap way ought to be found.

A promising solution suggested by Izadi et al. [26] may be a camera tracking certain recognizable shapes in front of the display.

3.3 Projection

Furthermore, there are various ways to achieve a large display. Many projects focus on projection-vision surfaces, but apply them differently. While these systems project an image on large context display, augmented reality applications walk the opposite way and project the enriched image on a focus display and have reality serve as context display.



Fig. 1. Occlusion and seamless integration using front projection [61].

One option is front projection, which means that a display (like a table or a wall) reflects the light/image of the projector into the eyes of the observers. A common setup for this scenario is a beamer projecting an image on a wall. The projector is usually mounted high on a shelf or on the ceiling to meet throw requirements and focal length of video cameras [9, 56, 61]. These mountings entail the drawbacks that they are difficult to setup and easily corrupted by minor vibrations. Additionally, once installed moving them becomes tedious [75]. While front projection does not need a cumbersome and possibly expensive setup with a diffusing layer and further devices in the back of the display, it fails to provide good results when the space before the display is crowded because the projection becomes occluded and thus badly visible. However, the occlusion/shadow can also be used as a depth cue and the handheld display can easier be integrated seamlessly with the projected image (see figure 1). To achieve such seamless integration while preventing unnecessary light of the projected image from hitting the handheld display, *Ubiquitous Graphics* [61] employs

a broadcast mechanism that updates every device when one device changes its location. Spatial awareness is provided by ultrasonic tracking of pens attached to the devices. Occluded areas are then masked so that the image is not projected in this specific area, but only in the free estate of the large display. This approach takes Baudisch's focus plus context screens [6] a step further. *PlayAnywhere* [75] takes a more compact and flexible approach to front projection. While most other applications rely on a fixed projection unit and separated devices for tracking and miscellaneous tasks, Wilson used a self-contained system comprised of a projector combined with an infrared camera and illuminant that can be placed directly on the table (see figure 2).



Fig. 2. Avoidance of occlusion by using a mobile projector [75].

The opposite, rear projection, is often used in table tops. It needs a diffusing layer on the surface that captures the projected light and thus makes it visible to the observer. Pioneering works concerning rear projection are *metaDESK* [69] and *HoloWall* [35]. The advantage of this approach is that it enables the construction of a self-contained system and easier positional tracking of the handheld devices as they can be traced on the usually touch-sensitive surface. It also eliminates problems evoked by occlusion. However, there are several drawbacks. One is the large amount of room space taken up by such a device. In contrast to front projection on walls which occupies almost no space, a dedicated surface usually included on a tabletop is necessary to operate a rear projection system. It is difficult to narrow down the space for the table as focal length and field of view of the camera has to be taken into consideration. Enabling the user to sit at the table may thus turn out to be difficult. Moreover, the imaging resolution is limited because the camera view is blurred by the diffuse surface [75].

A viable alternative to the common projection methods has been developed with *SecondLight* [26]. The system resembles a rear projection setup, but uses a switchable diffuser that enables and disables the diffusing layer in very fast alteration. In the enabled state the first image is projected onto the surface using common rear projection. When the diffusing layer is disabled, a second image is projected through the (now permeable) display. The first image remains visible to the human eye due to the imperceptibly fast switches. This enables a variety of new applications using the dimension above the display, such as thin sheets hold above it that provide a different view of the scene projected on the table surface (see figure 5).

3.4 Navigation

As the screen size of the handheld cannot display the whole background image and data at once, methods had to be found to cope with this problem. A well-known approach and the most widely applied method is *scrolling*. Scrolling features a static peephole with the content moving behind it. This technique benefits from the Parks Effect [48] which describes the storage of image data on the retina when only a small slit of it is visible at once. However, scrolling is counter-intuitive to the human mind as it defies spatial memory. The result is a loss of overview and inefficient navigation [20]. Mehra et al. [36] have proven that dynamic peephole navigation (moving the peephole over a static display) is superior to the common static method. Thus,

handhelds that are spatially aware and additionally capable of said dynamic interaction aid the human mind in navigation tasks. Rohs et al. [59] have further distinguished dynamic peephole navigation by focusing on whether visual context was provided or not. Their user study shows the higher efficiency of dynamic peephole navigation as a whole and only a very small difference concerning the existence of a visual context. This is probably due to the time-consuming switch of attention between the handheld and the static context on the large display.

If there is no visual context available, the question arises how you can locate off-screen objects. Panning is an option, but it is slow and tedious. Better methods are zooming, halo and the combination of both. Zooming means that the user can continually adjust the scale of the screen image and move to the desired object via *zoom pointing* [47] or *view pointing*. This operation obeys Fitts' Law [19]. Halo means that off-screen objects have a shiny circle that barely laps into the visible screen, thus indicating both direction and distance of the off-screen object. The drawback of the halo method is an overloaded screen when there are too many off-screen objects drawing their halo on the actual screen. The same argument applies for a similar technique called proxy that shows off-screen objects with a small representational icon at the margin of the visible screen. Zooming is limited by the resolution of the screen as distant objects become badly perceptible. Halo and zoom performed best in a user study conducted by Rohs et al. [58]. It is remarkable that the combination of both did not produce equal results, probably because users can only focus on one method at a time and thus have to decide which one to take. Zooming was also researched in a study of Olwal and Feiner [47] which featured rubbing, gesture, dragging and buttons as navigation techniques. While dragging resulted in high error rates, the other three techniques performed well.

Ballagas et al. [2] analyzed different techniques that can be applied to the interaction of handhelds and large displays. Positioning tasks can be managed as continuous indirect translation, for example, by a track pad which allows the user to move a cursor on the remote display (see [40]). A similar approach are directional step keys. Another option is a return-to-zero joystick that controls the velocity of the selected object or cursor. The operability of this concept has been proven in a study by Silfverberg et al. [65]. Positioning can also be done with an accelerometer (allowing navigation by tilting the device) [4, 52]. Tilt could alternatively be applied to further detail the input gained with other methods, thus easing and accelerating input [72]. Camera tracking is another option. It can be used in two different ways: either for continuous indirect translation like in *C-Blink* [38] or as direct translation, for example tracking an interactive light pen / laser pointer [1, 13, 29, 34, 45, 49]. One could also obtain discrete positioning by making a camera image and selecting the position wanted on the touch-sensitive screen of the handheld (used, for example, in *Point & Shoot* [3]). Adding a *sweep* mode that compares the relative motion of successive images leads to a continuous positioning again. This method also bears the advantage that the user does not have to point the handheld at the screen anymore nor does he have to watch the screen of the handheld. He can instead focus on the large display where the movable pointer resides and use his handheld as a *mouse-like* input device.

3D navigation was showcased by the *Boom Chameleon* [67], a successor of the pioneering Chameleon system [17], that used a video display mounted on a boom to show different views of a virtual scene depending on the position and orientation of the boom. A different, more immersive technique was used in *VITA* (Visual Interaction Tool for Archaeology) by Benko et. al [8] who visualized an archaeological excavation. The user can navigate in the virtual space of the site by gestures that are tracked via a glove and he can customize the view with speech commands.

3.5 Other forms of interaction

Besides navigational cues and positioning there are other forms of interaction that prove to be useful. These are orientation, selection, drag and drop, path, quantify and text.

3.5.1 Orientation

Many of the aforementioned techniques could also be used for orientation. A dedicated button could trigger an orientation mode that can then be operated using, for example, a velocity-controlled joystick or a track pad. *Point & Shoot* provides orientation by rotating the camera around the Z axis. (Augmented Reality) HUD (head-up display) applications use the orientation of the user himself/herself in the real world. The Boom Chameleon which can be oriented using its tracked boom can be seen as an example for such an application.

3.5.2 Selection

A virtual selection mechanism can be invoked by directly selecting the desired item in the real world. This is often made possible by tagging those items [70]. You could, for example, select a (tagged) product by taking it into your hands and/or placing it in a shopping cart and would then receive information on your tag-reading handheld device. That has been implemented in various projects like Xerox PARC's E-tag project [70], the Metro Future Store [33] and other projects harnessing RFID, NFC or visual tags to select items such as products [16, 23, 28]. RFID and NFC tags themselves have become inexpensive as of late, but need special components (RFID/NFC readers) to be read. These are often expensive and thus seldomly to be found in usual handhelds not specifically designed for this particular task.

Visual tags (Rekimoto's *CyberCode* [54]), in contrast, are cheap to produce and only need proper visual recognition software to be read. However, they bear the drawback that reading and processing the image data takes up a considerable amount of time because handheld devices usually still lack a floating point unit [11]. Additionally, few storage space is available compared to other tags and the readability of visual tags suffers from pollution.

In VITA the user can select items in 3D space by either grabbing it or pointing at them. Those selections can also be stored in a virtual tray for later use [8]. A different way is to use the camera of the handheld to directly pick an object on the large display by making an image of it [3, 12]. Objects can also be picked by a laser pointer that might be built into the handheld, followed by a fine-grained control of the selection on the handheld [39]. Another approach to selection is the use of speech recognition [44]. The user voices his commands, the handheld receives them using the built-in microphone, interprets them using pre-defined patterns and sends the coded commands to the large display to take effect.

Gesture recognition is another option to select an object. It can be implemented with any continuous positioning device such as velocity-controlled joysticks, accelerometers, camera tracked devices or track pads. Examples are the works of Patel et. al [49, 50] who used a handheld's accelerometric data to authenticate users wanting to view data on their mobile phone on a public terminal. A sequence of pre-defined movements was necessary to successfully authenticate. Another example is *VisionWand* by Cao and Balakrishnan [13] that features a non-electronic and un-tagged wand. Its postures and gestures are tracked to provide input such as selection. A more cumbersome and thus not frequently used method is typing the (previously assigned) character string name of the object on the handheld's keyboard. Another option that is not as cumbersome but inflexible is to assign buttons or functions to objects on a programmed function keyboard. The mechanism of *Point & Shoot* – selection via a camera image and an on-screen cursor – is more common and can also be implemented with different controls for the on-screen cursor, be it touch-screen entry, control through a track pad or other means [3].

3.5.3 Drag-And-Drop

Drag-and-drop combines both selection and positioning. On a usual tabletop, this action can be performed by pressing the finger on an item to select it and then making a continuous move to another location while keeping the finger on the surface. This is cumbersome and does not promote certain metaphors like cut & paste or a more flexible movement independent of the surface. Yee's *Sketchpad* [80] offers a prototypical 2D drag-and-drop feature for spatially aware handhelds

that can be activated by touching an object on the screen with the stylus. To move the object around, the user now has to move the handheld itself while still holding down onto the object. It is released again when the pen is lifted from the handheld display. A different approach is proposed by Sanneblad and Holmquist who implemented descriptive objects that can be left on a projected map and taken away again using the handheld device's GUI [61]. Other works detailing drag-and-drop techniques include Rekimoto's *HyperDrag* [56] and *Pick-and-Drop* [53] and Miller's and Myer's *Synchronized Clipboard* [37].

Nacenta et al. [41] offer a comparative study of possible drag-and-drop methods. These are *Pick-and-Drop* as implemented by Rekimoto, *Corresponding Gestures*, a similar approach that uses gestures instead of touch, *Slingshot*, *Pantograph* that employs forward movement in opposite to a slingshot, *Press-and-Flick* that features velocity calculated by pressure and *Radar View* that paints a little map for exact placement. The latter, which is suitable for spatially aware environments, performed best in all circumstances. For other environments Nacenta et al. suggest to use *Pick-and-Drop*. In his paper on *Peephole Displays*, Yee furthermore proposes a 3D drag-and-drop mechanism that is invoked by holding the pen on the object representation on the handheld screen (like in 2D mode) and then lifting the handheld away from the drawing plane upwards to a clipboard plane that is colored differently to provide an orientational cue. To paste the object back to the canvas plane again, the user holds the pen on the object in the clipboard and pushes his handheld down to the other plane. There (s)he may lift the pen and the item is released (see figure 3).

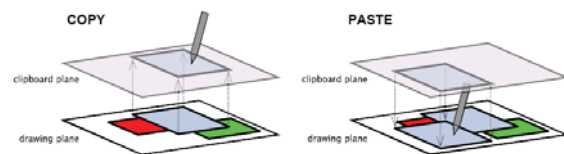


Fig. 3. Sketch for a 3D clipboard entailing drag-and-drop [80].

3.5.4 Pathing

Pathing means constituting a sequence of positions and orientations over time. While its task requirements differ from position and orientation, pathing nonetheless "adheres to the same taxonomy as the corresponding positioning and orienting techniques" [2]. Use cases of pathing include the arrangement of navigational instructions or the specification of a polygon to select an area.

3.5.5 Quantifying

Quantifying describes the input of a number within a certain range of numbers. This can easily be done with common numerical pads shipped with standard mobile phones. However, some handhelds do not offer such an input mechanism, so alternatives have to be found. These may be objects on the handheld screen that represent numbers (or rather a numerical pad) or tagged GUI elements that allow numeric input within a bounded range [57].

3.5.6 Text Input

Finally, text input may also be of concern in many environments. Most mobile phones offer text input using a mapping (for example, multi-press or T9) of the input using the numerical pad. Some even offer small-scale keyboards. The same could also be achieved by offering a touch-screen with the same approach. Developers have invested a lot of research in the field of mobile text input and there are numerous suggestions on how to enhance it, for example, by concurrent chording technique or using tilt [72, 73]. Other examples are key reassignment based on letter frequency [18, 43], improved methods to input foreign character sets [32, 31, 66] and adaptations such as track-ball entry or gestural input for impaired people [7, 77, 78, 79]. Text entry could also be managed by writing on the touch-sensitive screen of the handheld with a stylus (what could then in turn be displayed on the superimposed display such as a tabletop). As screen real estate

is small on handhelds, this method profits tremendously from spatial awareness, utilizing a considerably larger canvas by moving the device along while writing. This technique has been implemented in the *Peephole Doodle Pad* [80]. A similar approach is taken in *Ubiquitous Graphics* where one can add annotations to a projected map by writing on the handheld display [61].

Ballagas et al. [3] also mention the scarce use of acoustic and haptic feedback which might prove as a field of future research.

4 THREE DIFFERENT TYPES OF SUPERIMPOSED DISPLAYS

To showcase the most salient works and thereby provide an impression of possible applications, the subject shall now be partitioned into three parts. The portraits may also serve to deepen the specific techniques and technologies applied in the different fields. Please bear in mind that any criticism shall not detract from the achievements of the projects presented, but only exemplify improvement opportunities and exhibit alternatives to the technologies used. First the paper will expose applications based on touch, then provide insight into past and recent research addressing interaction beyond the surface and in the last section it will show applications for superimposed displays in the domain of augmented reality.

4.1 Touch-Based Interaction

The item most widely covered by research is touch-based interaction. The user has to place the superimposed display on the surface to have it recognized, tracked and updated with content. Here, *LightSense* by Alex Olwal [46] shall receive an exemplary review.

The system has been dubbed after the gradually distributing ultra bright LEDs in modern mobile phones equipped with a digital camera that are employed as readily identifiable active markers. Unlike other rear projection tabletops this system is not self-contained but is comprised of a laptop with attached camera and a diffusing surface. The second outstanding difference is the use of the diffusing surface – its purpose is not to display the projection, but to aid the proper monitoring of the handheld on the surface. The system’s primary goal is to provide a solid framework to enable spatial awareness for common handheld devices and thereby promote interactivity.

Unfortunately, the paper fails to describe the connection process in depth. It can be inferred though that a connection is initiated by placing the device on the surface. The system features a Bluetooth server that subsequently connects to the handheld on the surface. This process might stir up problems such as ambiguity if more than one device is in range. So far, *LightSense* does not provide a solution for this issue, yet Olwal mentions that tracking of multiple devices and disambiguation might be introduced by time division multiplexing and controlling the LED of the device. This suggestion reminds of *BlueTable’s* vision-based handshaking. Currently most handhelds do not allow programmable control of the LED which is detrimental to this solution if the system is aimed to include as much handheld devices as possible.

The handheld is tracked by locating its light source on the diffusing surface (see figure 4). The aforementioned camera monitors the surface from the back. Its images are filtered to a binary state and otherwise processed to exclude noisy regions. The algorithm then tries to fit an ellipsoid around the remaining pixels in the resulting image whose center marks the position of the handheld. The distance from the surface may also be inferred by adapting the filters to more blurry lightspots indicating a light source further away from the surface. Olwal achieved a discrete ten-grade scale to measure depth, so continuous tracking only works in two dimensions. This might actually classify the project at hand as an application harnessing the dimension above the surface (and it de facto does that concerning zooming). However, such a discrete scale judged by luminance already prompts image processing algorithms with problems as disambiguation reaches its limits. Further confining obstacles in that respect accrue from ambient light. Olwal tries to circumvent this issue by providing manual calibration and tweaking at runtime. *LightSense* also employs a second means to track the device: ubiquitous sensing tags. It does not rely

on RFID or NFC for that purpose (as did, for example, *Touch & Interact* by Hardy and Rukzio [23]), but adopts Light Dependent Resistors (LDRs) interfaced with a microcontroller. Olwal’s fond criticism of the radio-based technologies (large tag size, discrete nature, inconsistencies due to overlapping and sensitivity to surrounding tags) might best be comprehensible observing that LDRs come in handy for the specific setup with light recognition, but they are probably not suitable for other tracking options.

There is no projection present in *LightSense*. A physical subway map is used to showcase the technology. When the user hovers the handheld over a station, its screen shows the station’s periphery. The handheld thus serves as a peephole and according navigation is provided. The zoom feature proposed by Yee [80] – zooming by changing the distance between the handheld and the surface – is implemented additionally. It is also remarkable that audio and vibrotactile feedback are integrated to playback the names of stations or indicate subway lines. This can help impaired users tremendously – a group of customers that often goes unnoticed by other projects.

Regarding other forms of interaction aside of navigation, the paper proposes to use multiple cameras to detect orientation judged by light intensity. However, the devices would need an elongated light source or more than one LED to detect their orientation on the surface plane by shape analysis. Tilt could be detected by using the handheld’s own camera to look for reflections of its light.

Selection is not genuinely supported. It may be deemed sufficient to “select” a station by hovering over it, but more sophisticated implementations cannot be found. Still, they are imaginable, for example, a user could be enabled to select stations to get information for the trip such as costs or even the ticket. This would also be a possible implementation of pathing which is equally not offered. Another example would be to view the adequate subway timetable. Moreover, drag and drop is not featured. This is not surprising as the surface remains static and only the handheld shows interactive content (reminiscent of the *focus + context* metaphor [6] mentioned earlier). Other interactive techniques such as quantify and text are also not provided. An option to enrich the application would have been to allow annotations for the subway stations or surroundings as it is supported, for instance, by *Ubiquitous Graphics* [61].



Fig. 4. Tracking of the handheld using its LED [46].

4.2 Beyond The Surface

Several projects have attempted to get rid of the constraints of touch-based interaction which narrows interaction to the flat surface. The impetus therefore was to conquer the dimension above the surface. Reliable techniques are still scarce, but some ways to overcome the former constraints have been shown in a few recent papers. This section shall detail the ambitious *SecondLight* project [26].

This work received by Izadi et al. its name from the switchable diffuser that once projects an image on the diffusing surface and is alternately permeated by a second image by shortly switching off the diffusing layer. The second image may then be captured by layers or other diffuse objects above the surface. The diffuser itself is comprised of electronically controllable liquid crystal material that allows rapid switching as fast as necessary to prevent flickering. Two independent

projection channels can be thus be harnessed properly. This allows for applications that benefit from spatial awareness and magic lenses as envisioned by Bier et al. [10]. *SecondLight* provides an example where a superimposed translucent sheet of diffuse film unveils the inwards of a car shown on the context display (see figure 5). Another example is a star map on the tabletop and the mobile surface showing constellations in the superimposed area.

Connecting to devices above the display is not genuinely supported. Tracking, however, is a strong point of the system. Adopting Han’s pioneering work [22], it employs frustrated total internal reflection (FTIR) to detect fingertips or others objects touching the surface. Additionally, IR reflective objects can be detected in the clear phase using diffuse illumination. Visual markers or sophisticated recognition algorithms may be an option here. Diffuse illumination is also a solid way to infer depth cues as objects further away from the surface become blurred gradually. Layers like the depicted film in the exemplary figure (5) are tracked by monitoring four distinctive points on the layer. This can be achieved with diffuse IR illumination and passive markers reflecting the infrared light. Izadi et al. used active markers instead as the aforementioned technique produces noisy results.

SecondLight employs rear projection for both images enabling a self-contained system. Two projectors are used that operate at 60 Hz and are blocked alternately by optical shutters that utilize the same technology as the switchable diffuser (ferroelectric liquid crystal). This entails the unique feature of projecting an image beyond the surface onto a mobile layer. This layer may also be moved, tilted or otherwise modified and the contained image remains the same since projective transformation is applied.

The system offers navigation like other peephole-based projects: the image is continuously “updated” on the mobile display while it resides in the area above the context display. As the second image is always projected into the space above the surface, spatial awareness of the handheld is naturally inferred and does not have to be computed which greatly reduces complexity. Furthermore, navigation may also happen on the mobile surfaces as multi-touch input on them can be sensed using IR. Orientation is detected by an inside-out approach, namely monitoring the four edges of the mobile surface. While this may seem like a simple approach, one has to bear in mind that each handheld device has to be equipped with markers. This adds complexity and especially inflexibility to the system as translucent layers cannot be used to full extent on the fly. Selections on the tabletop surface are monitored by a common multi-touch detection algorithm. On the mobile surfaces they can be detected by the already mentioned multi-touch input. This is an interesting feature considering that works that enable distant selections are still scarce. The system offers zooming by changing the distance between mobile layer and surface, a technology we already witnessed in *LightSense* [46] and Yee’s projects [80]. Drag and drop, however, is not supported. The same applies for pathing, quantify and text. Aside from gesture-based input adopted from *TouchLight* [74], this lack of interactive features may be a source for future work based on *SecondLight* or its proposed technology.



Fig. 5. Beyond-the-surface projection with a switchable diffuser [26].

4.3 Augmented Reality

While both touch-based interaction and beyond-the-surface interaction mostly happen on either tabletops or walls superimposed by a hand-held device, augmented reality follows a different paradigm: the real world serves as the large display that is visually enriched by HUD applications. The *Simulated Augmented Reality Windshield Display* by Kim and Dey [27] shall serve as the example of choice.

As the name indicates, the prototype only features a simulated environment. Scientific research papers on this topic are hard to find. Sato et al. [62] is probably the most similar realization; other works paved the way for the project at hand. Aside of that, several commercial proof-of-concept implementations have already been tested¹. Use cases are viewing car-based information such as speed or fuel level, surroundings such as sights or petrol stations or providing navigational cues and aid for elder drivers as envisioned by the *Simulated AR Windshield*.

Applying the theory mentioned above, first one notices first that augmented reality applications demand different views on connection and tracking. This paper will not dwell on “tracking” reality, but only refer to sources that cover the subject in depth [11, 25, 42, 64]. While other projects use a head-mounted display (for example, VITA [8] or *Wearable Virtual Tablet* [68]), this application projects its images onto the windshield of a car. This method entails the advantage that no bothersome device has to be worn while information is still integrated seamlessly into the user’s field of view. However, it also evokes concerns if this projection might not distract the driver from his cardinal tasks. These objections are soothed by user study conducted by Kim and Dey which provides the result that in fact the system reduces divided attention because the driver does not have to change his focus between interior displays and the street anymore. It is reported though that some users interpreted the system’s highlighted route of navigation as a admission to cross intersection regardless of potentially prohibiting traffic lights. This observation may incite future developments that allow communication with infrastructure such as said traffic lights. The paper does not feature a specific projection method, but only assumes that the particular problems arising from projection on curved glass planes and intertwining of reality and windshield display “will be addressed in the near future by other researchers” [27]. Concerning the projected image, it proposes a 2D bird-view image of the roads ahead that continually slides down into view as the car progresses (see figure 6).

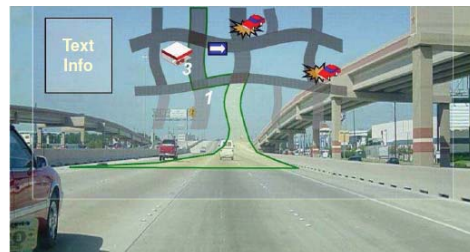


Fig. 6. AR windshield displaying a map to aid navigation [27].

Absolute positional data for this case can be acquired via GPS or similar location-based services. This can be deemed enough for the intended purpose concerning spatial resolution. Orientation may be inferred either passively by direction of movement – which is the common way used in today’s car navigation systems – or actively by in-car-monitoring of the movement. Selecting something on the screen is not featured in this project. This is not surprising as the windshield shall only serve as a passive screen showing information and not animate the driver to be distracted by interactivity. In other circumstances, however, such a feature is likely to prove useful. Implementations can be

¹<http://www.youtube.com/watch?v=kzLK6UnQcnU>
<http://www.mvs.net/>
<http://www.youtube.com/watch?v=tA8VZUqo6FQ>

observed in the aforementioned papers on augmented reality projects. The same applies for drag and drop as well as quantify. Pathing is obviously a strong point of the system and is inferred by the GPS data. An interesting additional feature would be interactive route planning on the windshield screen, but input may emerge as a obstruction in this regard. Moreover, text entry may be enhanced with little effort. Given an interactive screen, it is perfectly feasible to enter text, for instance, to remember taverns offering delicious food. Interactivity should, however, be limited to times when the engine is stopped.

Only this specific paper has been reviewed as delving into augmented reality as a whole would exceed this paper by far. For a comprehensive aggregation of research please refer to Bimber and Raskar's *Modern approaches to augmented reality* [11].

5 CONCLUSION

A profound overview has been given of the different fields where superimposed displays can be applied. It is apparent that these are touch-based interaction, use of the dimension beyond the surface and applications in the domain of augmented reality. The technologies necessary to achieve superimposed views have been detailed. There are the issues of connecting the devices (...) and tracking the mobile hand-helds (...). The process of projecting an image on the large display and mapping the large image onto the small display has been depicted. A promising approach in this matter is the introduction of a switchable diffuser that facilitates the use of the dimension above the surface to display additional information. Different options to enable navigation as well as other interactive tasks such as selection or orientation have been presented. Furthermore, insight has been supplied on what difficulties are still there for developers. Several constraints were mentioned that yet prevent a breakthrough of the aforementioned technologies to everyday life. Impediments that need to be overcome remain especially in the area of connection, smooth tracking and rich and coherent interaction.

With the dissemination of mobile phones, a new age of ubiquitous computing is dawning. This prospect bears countless possibilities for new innovative applications in this domain. This paper has hopefully provided a starting point for developers looking for preceding and compelling works in this thriving new field of research.

REFERENCES

- [1] B. A. Ahlborn, D. Thompson, O. Kreylos, B. Hamann, and O. G. Staadt. A practical system for laser pointer interaction on large displays. In *VRST '05*, pages 106–109, New York, NY, USA, 2005. ACM.
- [2] R. Ballagas, J. Borchers, M. Rohs, and J. G. Sheridan. The smart phone: A ubiquitous input device. *IEEE Pervasive Computing*, 5(1):70, 2006.
- [3] R. Ballagas, M. Rohs, and J. G. Sheridan. Sweep and point and shoot: phonecam-based interactions for large public displays. In *CHI '05*, pages 1200–1203, New York, NY, USA, 2005. ACM.
- [4] J. F. Bartlett. Rock 'n' scroll is here to stay. *IEEE Comput. Graph. Appl.*, 20(3):40–45, 2000.
- [5] J. J. Barton, S. Nylander, F. Folowosele, and B. Harrison. Dialing for displays: Session initiation protocol for opportunistic augmentation. In *PERCOM '06*, pages 60–65, Washington, DC, USA, 2006. IEEE Computer Society.
- [6] P. Baudisch, N. Good, and P. Stewart. Focus plus context screens: combining display technology with visualization techniques. In *UIST '01*, pages 31–40, New York, NY, USA, 2001. ACM.
- [7] M. Belatar and F. Poirier. Text entry for mobile devices and users with severe motor impairments: handiglyph, a primitive shapes based onscreen keyboard. In *Assets '08*, pages 209–216, New York, NY, USA, 2008. ACM.
- [8] H. Benko, E. W. Ishak, and S. Feiner. Collaborative mixed reality visualization of an archaeological excavation. In *ISMAR '04*, pages 132–140, Washington, DC, USA, 2004. IEEE Computer Society.
- [9] F. Berard. The magic table: Computer vision based augmentation of a whiteboard for creative meetings. In *Proceedings of the IEEE International Conference in Computer Vision*, 2003.
- [10] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose. Toolglass and magic lenses: the see-through interface. In *SIGGRAPH '93*, pages 73–80, New York, NY, USA, 1993. ACM.
- [11] O. Bimber and R. Raskar. Modern approaches to augmented reality. In *ACM SIGGRAPH '07 Courses*, page 1, New York, NY, USA, 2007. ACM.
- [12] S. Boring, M. Altendorfer, G. Broll, O. Hilliges, and A. Butz. Shoot & copy: phonecam-based information transfer from public displays onto mobile phones. In *Mobility '07*, pages 24–31, New York, NY, USA, 2007. ACM.
- [13] X. Cao and R. Balakrishnan. Visionwand: interaction techniques for large displays using a passive wand tracked in 3d. In *UIST '03*, pages 173–182, New York, NY, USA, 2003. ACM.
- [14] K. C. Cheung, S. S. Intille, and K. Larson. An inexpensive bluetooth-based indoor positioning hack. In *In Ubicomp Poster Proceedings*, 2006.
- [15] M. D. Corner and B. D. Noble. Zero-interaction authentication. In *MobiCom '02*, pages 1–11, New York, NY, USA, 2002. ACM.
- [16] K. Finkenzeller. *RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification*. John Wiley & Sons, Inc., New York, NY, USA, 2003.
- [17] G. W. Fitzmaurice. Situated information spaces and spatially aware palm-top computers. *Commun. ACM*, 36(7):39–49, 1993.
- [18] N. Green, J. Kruger, C. Faldu, and R. St. Amant. A reduced qwerty keyboard for mobile text entry. In *CHI '04*, pages 1429–1432, New York, NY, USA, 2004. ACM.
- [19] Y. Guiard and M. Beaudouin-Lafon. Target acquisition in multiscale electronic worlds. *Int. J. Hum.-Comput. Stud.*, 61(6):875–905, 2004.
- [20] Y. Guiard, M. Beaudouin-Lafon, J. Bastin, D. Pasveer, and S. Zhai. View size and pointing difficulty in multi-scale navigation. In *AVI '04*, pages 117–124, New York, NY, USA, 2004. ACM.
- [21] M. Haase and M. Handy. Bluetrack – imperceptible tracking of bluetooth devices. In *In Ubicomp Poster Proceedings*, 2004.
- [22] J. Y. Han. Low-cost multi-touch sensing through frustrated total internal reflection. In *UIST '05*, pages 115–118, New York, NY, USA, 2005. ACM.
- [23] R. Hardy and E. Rukzio. Touch & interact: touch-based interaction of mobile phones with displays. In *MobileHCI '08*, pages 245–254, New York, NY, USA, 2008. ACM.
- [24] L. E. Holmquist, F. Mattern, B. Schiele, P. Alahuhta, M. Beigl, and H.-W. Gellersen. Smart-its friends: A technique for users to easily establish connections between smart artefacts. In *UbiComp '01*, pages 116–122, London, UK, 2001. Springer-Verlag.
- [25] Z. Hu and K. Uchimura. Solution of camera registration problem via 3d-2d parameterized model matching for on-road navigation. *Intl Journal of Image and Graphics*, 4(1):3–20, 2004.
- [26] S. Izadi, S. Hodges, S. Taylor, D. Rosenfeld, N. Villar, A. Butler, and J. Westhues. Going beyond the display: a surface technology with an electronically switchable diffuser. In *UIST '08*, pages 269–278, New York, NY, USA, 2008. ACM.
- [27] S. Kim and A. K. Dey. Simulated augmented reality windshield display as a cognitive mapping aid for elder driver navigation. In *CHI '09*, pages 133–142, New York, NY, USA, 2009. ACM.
- [28] S. Konomi and G. Roussos. Ubiquitous computing in the real world: lessons learnt from large scale rfid deployments. *Personal Ubiquitous Comput.*, 11(7):507–521, 2007.
- [29] D. Kurz, F. Hantsch, M. Grobe, A. Schiewe, and O. Bimber. Laser pointer tracking in projector-augmented architectural environments. In *ISMAR '07*, pages 1–8, Washington, DC, USA, 2007. IEEE Computer Society.
- [30] J. Lester, B. Hannaford, and G. Borriello. Are you with me? – using accelerometers to determine if two devices are carried by the same person. In *2nd International Conference on Pervasive Computing*, 2004.
- [31] Y. Liu and K.-J. Rähkä. Rotatxt: Chinese pinyin input with a rotator. In *MobileHCI '08*, pages 225–233, New York, NY, USA, 2008. ACM.
- [32] Y. Liu and Q. Wang. Chinese pinyin phrasal input on mobile phone: usability and developing trends. In *Mobility '07*, pages 540–546, New York, NY, USA, 2007. ACM.
- [33] C. Loebbecke. Rfid technology and applications in the retail supply chain: The early metro group pilot. In *18th Bled eConference – eIntegration in Action*, Kranj, Slovenia, 2005. University of Maribor.
- [34] A. Madhavapeddy, D. Scott, R. Sharp, and E. Upton. Using camera-phones to enhance human-computer interaction. In *UbiComp '04 (Adjunct Proceedings: Demos)*, 2004.
- [35] N. Matsushita and J. Rekimoto. Holowall: designing a finger, hand, body, and object sensitive wall. In *UIST '97*, pages 209–210, New York, NY, USA, 1997. ACM.
- [36] S. Mehra, P. Werkhoven, and M. Worring. Navigating on handheld dis-

- plays: Dynamic versus static peephole navigation. *ACM Trans. Comput.-Hum. Interact.*, 13(4):448–457, 2006.
- [37] R. C. Miller and B. A. Myers. Synchronizing clipboards of multiple computers. In *UIST '99*, pages 65–66, New York, NY, USA, 1999. ACM.
- [38] K. Miyaoku, S. Higashino, and Y. Tonomura. C-blink: a hue-difference-based light signal marker for large screen interaction via any mobile terminal. In *UIST '04*, pages 147–156, New York, NY, USA, 2004. ACM.
- [39] B. A. Myers, R. Bhatnagar, J. Nichols, C. H. Peck, D. Kong, R. Miller, and A. C. Long. Interacting at a distance: measuring the performance of laser pointers and other devices. In *CHI '02*, pages 33–40, New York, NY, USA, 2002. ACM.
- [40] B. A. Myers, H. Stiel, and R. Gargiulo. Collaboration using multiple pdas connected to a pc. In *CSCW '98*, pages 285–294, New York, NY, USA, 1998. ACM.
- [41] M. A. Nacenta, D. Aliakseyeu, S. Subramanian, and C. Gutwin. A comparison of techniques for multi-display reaching. In *CHI '05*, pages 371–380, New York, NY, USA, 2005. ACM.
- [42] W. Narzt, G. Pomberger, A. Ferscha, D. Kolb, R. Muller, J. Wieghardt, H. Hortner, and C. Lindinger. Augmented reality navigation systems. *Universal Access Information Society*, 4(3):177–187, 2006.
- [43] S. B. Nesbat. A system for fast, full-text entry for small electronic devices. In *ICMI '03*, pages 4–11, New York, NY, USA, 2003. ACM.
- [44] J. Nichols, B. A. Myers, M. Higgins, J. Hughes, T. K. Harris, R. Rosenfeld, and K. Litwack. Personal universal controllers: controlling complex appliances with guis and speech. In *CHI '03*, pages 624–625, New York, NY, USA, 2003. ACM.
- [45] D. R. Olsen, Jr. and T. Nielsen. Laser pointer interaction. In *CHI '01*, pages 17–22, New York, NY, USA, 2001. ACM.
- [46] A. Olwal. Lightsense: enabling spatially aware handheld interaction devices. In *ISMAR '06*, pages 119–122, Washington, DC, USA, 2006. IEEE Computer Society.
- [47] A. Olwal and S. Feiner. Spatially aware handhelds for high-precision tangible interaction with large displays. In *TEI '09*, pages 181–188, New York, NY, USA, 2009. ACM.
- [48] T. Parks. Post-retinal visual storage. *American Journal of Psychology*, 78(1):145–147, 1965.
- [49] S. N. Patel and G. D. Abowd. A 2-way laser-assisted selection scheme for handhelds in a physical environment. In A. K. Dey, A. Schmidt, and J. F. McCarthy, editors, *UbiComp '03*, volume 2864 of *Lecture Notes in Computer Science*, pages 200–207. Springer, 2003.
- [50] S. N. Patel, J. S. Pierce, and G. D. Abowd. A gesture-based authentication scheme for untrusted public terminals. In *UIST '04*, pages 157–160, New York, NY, USA, 2004. ACM.
- [51] T. Pering, R. Ballagas, and R. Want. Spontaneous marriages of mobile devices and interactive spaces. *Commun. ACM*, 48(9):53–59, 2005.
- [52] M. Rahman, S. Gustafson, P. Irani, and S. Subramanian. Tilt techniques: investigating the dexterity of wrist-based input. In *CHI '09*, pages 1943–1952, New York, NY, USA, 2009. ACM.
- [53] J. Rekimoto. Pick-and-drop: a direct manipulation technique for multiple computer environments. In *UIST '97*, pages 31–39, New York, NY, USA, 1997. ACM.
- [54] J. Rekimoto and Y. Ayatsuka. Cybercode: designing augmented reality environments with visual tags. In *DARE '00*, pages 1–10, New York, NY, USA, 2000. ACM.
- [55] J. Rekimoto, Y. Ayatsuka, and M. Kohno. Synctap: An interaction technique for mobile networking. In *Mobile HCI '03*, pages 104–115, Berlin / Heidelberg, 2003. Springer.
- [56] J. Rekimoto and M. Saitoh. Augmented surfaces: a spatially continuous work space for hybrid computing environments. In *CHI '99*, pages 378–385, New York, NY, USA, 1999. ACM.
- [57] M. Rohs. Visual code widgets for marker-based interaction. In *ICD-CSW '05*, pages 506–513, Washington, DC, USA, 2005. IEEE Computer Society.
- [58] M. Rohs and G. Essl. Which one is better?: information navigation techniques for spatially aware handheld displays. In *ICMI '06*, pages 100–107, New York, NY, USA, 2006. ACM.
- [59] M. Rohs, J. Schöning, M. Raubal, G. Essl, and A. Krüger. Map navigation with mobile devices: virtual versus physical movement with and without visual context. In *ICMI '07*, pages 146–153, New York, NY, USA, 2007. ACM.
- [60] E. Rukzio, K. Leichtenstern, V. Callaghan, A. Schmidt, P. Holleis, and J. Chin. An experimental comparison of physical mobile interaction techniques: Touching, pointing and scanning. In *UbiComp '06*, pages 87–104, Berlin / Heidelberg, 2006. Springer.
- [61] J. Sanneblad and L. E. Holmquist. Ubiquitous graphics: combining handheld and wall-size displays to interact with large images. In *AVI '06*, pages 373–377, New York, NY, USA, 2006. ACM.
- [62] A. Sato, I. Kitahara, Y. Kameda, and Y. Ohta. Visual navigation system on windshield head-up display. In *Proceedings of 13th World Congress on Intelligent Transport Systems*, 2006.
- [63] D. Scott, R. Sharp, A. Madhavapeddy, and E. Upton. Using visual tags to bypass bluetooth device discovery. *SIGMOBILE Mob. Comput. Commun. Rev.*, 9(1):41–53, 2005.
- [64] S. Scott-Young. Seeing the road ahead: Gps-augmented reality aids drivers. *GPS World*, 14(11):22–28, 2003.
- [65] M. Silfverberg, I. S. MacKenzie, and T. Kauppinen. An isometric joystick as a pointing device for handheld information terminals. In *GI '01*, pages 119–126, June 2001.
- [66] Y. K. Thu and Y. Urano. Positional mapping: keyboard mapping based on characters writing positions for mobile devices. In *ICMI '07*, pages 110–117, New York, NY, USA, 2007. ACM.
- [67] M. Tsang, G. W. Fitzmaurice, G. Kurtenbach, A. Khan, and B. Buxton. Boom chameleon: simultaneous capture of 3d viewpoint, voice and gesture annotations on a spatially-aware display. In *UIST '02*, pages 111–120, New York, NY, USA, 2002. ACM.
- [68] N. Ukita and M. Kidode. Wearable virtual tablet: fingertip drawing on a portable plane-object using an active-infrared camera. In *IUI '04*, pages 169–176, New York, NY, USA, 2004. ACM.
- [69] B. Ullmer and H. Ishii. The metadesk: models and prototypes for tangible user interfaces. In *UIST '97*, pages 223–232, New York, NY, USA, 1997. ACM.
- [70] R. Want, K. P. Fishkin, A. Gujar, and B. L. Harrison. Bridging physical and virtual worlds with electronic tags. In *CHI '99*, pages 370–377, New York, NY, USA, 1999. ACM.
- [71] M. Weiser. Some computer science issues in ubiquitous computing. *Commun. ACM*, 36(7):75–84, 1993.
- [72] D. Wigdor and R. Balakrishnan. Tilttext: using tilt for text input to mobile phones. In *UIST '03*, pages 81–90, New York, NY, USA, 2003. ACM.
- [73] D. Wigdor and R. Balakrishnan. A comparison of consecutive and concurrent input text entry techniques for mobile phones. In *CHI '04*, pages 81–88, New York, NY, USA, 2004. ACM.
- [74] A. D. Wilson. Touchlight: an imaging touch screen and display for gesture-based interaction. In *ICMI '04*, pages 69–76, New York, NY, USA, 2004. ACM.
- [75] A. D. Wilson. Playanywhere: a compact interactive tabletop projection-vision system. In *UIST '05*, pages 83–92, New York, NY, USA, 2005. ACM.
- [76] A. D. Wilson and R. Sarin. Bluetable: connecting wireless mobile devices on interactive surfaces using vision-based handshaking. In *GI '07*, pages 119–125, New York, NY, USA, 2007. ACM.
- [77] J. Wobbrock and B. Myers. Trackball text entry for people with motor impairments. In *CHI '06*, pages 479–488, New York, NY, USA, 2006. ACM.
- [78] J. O. Wobbrock, D. H. Chau, and B. A. Myers. An alternative to push, press, and tap-tap-tap: gesturing on an isometric joystick for mobile phone text entry. In *CHI '07*, pages 667–676, New York, NY, USA, 2007. ACM.
- [79] J. O. Wobbrock and B. A. Myers. Gestural text entry on multiple devices. In *Assets '05*, pages 184–185, New York, NY, USA, 2005. ACM.
- [80] K.-P. Yee. Peephole displays: pen interaction on spatially aware handheld computers. In *CHI '03*, pages 1–8, New York, NY, USA, 2003. ACM.

Superimposed Displays

Sascha Gebhardt

Abstract— This paper gives a short overview over the current state of science around the field of superimposed displays. There will be two major sections. The first one is presenting technologies that can be used to realize spatial awareness, what is a very important feature around the superimposing of displays. The second one is presenting various possibilities of superimposing displays, realized by various different scientists. It is split up into three main sections: virtual superimposing, distant superimposing and direct superimposing.

Index Terms—superimposed displays, tabletop, wall-display, handheld, multi-touch, tracking, spatial awareness

1 INTRODUCTION

Several different kinds of displays are used today, from small mobile phone displays, up to very large wall-sized displays everything is possible. There are quite a lot technologies, that can be used to realize displays, like CRT-, LCD-, Plasma-, OLED-screens or front- or rear-projecting LCD- or DLP-projectors. All of these different display variants have advantages, due to whom they have the right to exist. These advantages can be low production costs, robustness, longevity, low energy consumption, high contrast or brightness, high viewing angles, low weight, high resolution and so on. By combining different kinds of displays, or multiple displays of the same kind, some of these advantages can be enhanced or new advantages can be created. So for example two LCDs can be placed next to each other to gain a higher desktop resolution, if the corresponding computer supports it.

Another possibility of the combination of displays is superimposing them. Superimposed displays can be useful for various purposes: realizing focus and context techniques, augmentation, handling large amounts of complex data or enhancing the local or overall display resolution. This paper is about to give an overview over the current state of science in the field of superimposed displays.

It will be split up into two major parts: at first it will shortly present a selection of different tracking technologies, that enable spatial awareness, what often is an essential feature to realize superimposed displays. After that the main part will follow, which will, one after another, cover the three main categories amongst superimposed displays, which are virtual superimposing, distant superimposing and direct superimposing of displays. Finally a short conclusion will be given.

2 TRACKING TECHNOLOGIES

An important feature for the realization of superimposing physical displays is spatial awareness. That means, that the relative positions of the superimposed displays to each other must be known. This section will give a short introduction on how this spatial awareness can be realized. The first solution gains the spatial awareness of a mobile phone by optically tracking the positions of its photo light. The next technique is to track a display on top of another with a camera that is mounted in front of both of the displays. After that we will present a technique of tracking a device on top of a multi-touch table by generating a unique pattern of touch-points. Finally we will have a look on tracking by applying and identifying objects with the use of markers.

2.1 Tracking with the Photo Light of a Mobile Phone

Alex Olwal presented its LightSense [14] technology. It realizes spatial awareness of a mobile phone, by enabling its photo light and opti-

cally tracking the position of the light blob. He presents two different solutions on how this tracking can be realized.

The first one is continuous tracking for tables and walls. Here he places a camera behind a diffusion surface (like a multi-touch table) and tracks the light blob that is generated by the photo light with a camera by applying several image filters like thresholding, noise reduction and fitting. In the result he detects an ellipse, whose center is the actual position of the photo light. Additionally, by analyzing the intensity and size of this ellipse, information about the distance of the mobile phone to the surface can be gathered, though only discrete values are obtained.

The second solution is inspired by Radio Frequency Identification (RFID) Near Field Communication (NFC) tags and is meant for applications, that do not benefit from continuous tracking: the so called hotspot-tracking. To realize this, Olwal placed several Light Dependent Resistors (LDRs) in strategic positions behind a diffusion surface. The LDRs are connected to a micro controller and their voltage level is delivered to the mobile phone by Bluetooth. The mobile phone can then calculate its position above the surface.

An advantage of the solutions in LightSense [14] is, that no hardware modifications need to be done to the mobile phone and that the second solution can be used with rather thin surfaces, without any visible tracking technology. On the other hand, missing orientation information are a rather large handicap.

2.2 Tracking with a Front-Mounted Camera

Andrew Wilson presented PlayAnywhere [17] and based on it BlueTable [18], who both use tracking with a camera that is mounted in front of a front-projected surface. In PlayAnywhere he presents several tracking and recognition algorithms that apply to various different demands. A shadow-based touch detection algorithm was implemented as well as a fast, simple visual bar code scheme, the ability to track sheets of paper and an optical flow-based algorithm for the manipulation of onscreen objects that does not rely on fragile tracking algorithms. The bar code and paper sheet trackings can be used to realize virtual displays on top of the front-projected surface.

In BlueTable [18] he presents a vision-based handshaking process between the front-projected surface already presented in PlayAnywhere and a mobile phone. The spatial awareness of the phone is realized by visually recognizing phone-like objects and identifying them with the aid of the Bluetooth and Infrared Data Association (IrDA) communication ports or the mobile's display. The BlueTable asks the phone to blink its IrDA port or its display with a specific code. This blinking is then recognized by the camera, and the handshaking-process is finished. From now on the phone-like shape is associated with a specific phone and spatial awareness is established.

A quite nice feature of PlayAnywhere is, that the projection and tracking hardware is combined in one mobile unit. As well PlayAnywhere as BlueTable offer 2D tracking with orientation information but are vulnerable to objects blocking the line-of-sight from the camera to the tracked object.

-
- Sascha Gebhardt is studying Media Informatics at the University of Munich, Germany, E-mail: sascha.gebhardt@t-online.de
 - This research paper was written for the Media Informatics Advanced Seminar on Interactive Surfaces, 2009

2.3 Tracking by Generating a Pattern of Touch Points

In iPodParty [8] a tracking approach was presented, that is based on the generation of a unique pattern of touch points. Tracking carriages were mounted to iPod touches, that generated unique patterns of three touch points on the surface of a Frustrated Total Internal Reflection (FTIR) [9] multi-touch table. Two different kinds of carriages were invented: a heavy weighted one which generated touch points by pressure and an electrical one which sent infrared light through the surface of the multi-touch table to generate the touch points. With these patterns an iPod touch could be identified distinctly and its position and orientation to the table could be calculated.

Benefits from this approach are, that the tracking itself is quite cheap on the computational side, though the general touch recognition has computation costs. The need to mount rather big tracking carriages is a disadvantage on the optical and haptic side.

2.4 Tracking Objects with Markers

Another approach of tracking objects is to place visual markers on them, to identify and track them. Markers usually are unique black and white or colored patterns. Examples for projects/papers concerning marker tracking are [6, 17, 2, 11, 13, 1]. Problems with markers are, that they usually do not look very nice and that they need a quite constant light as well as a sufficient optical resolution to work properly. But if acceptable light conditions are present and the tracking optics have enough resolution and are near enough, they make a reliable and robust tracking possible.

2.5 Comparison of the Different Tracking Technologies

All of the technologies that were presented in this section work basically, but all of them have their very own advantages and disadvantages. The quality of a tracking system can be judged in four main categories: robustness to external influences, tracking dimensions, the computation needs as well as the hardware modifications, that need to be done.

All of the tracking technologies presented here are robust under a laboratory environment, but due to the fact that all of them are optically based systems, extreme light variations could probably harm their robustness. For example iPod Party [8] would hardly work when placed outside on a sunny day where clouds pass the sun, because these light variations would be simply too intense for the tracking. PlayAnywhere [17] and BlueTable [18] are also sensitive to occlusion of the objects to track, because of the front mounted camera, that needs line-of-sight to track objects.

There are seven possible dimensions of tracking: 2D-position on a surface, distance and 1D-orientation on a surface or 3D-orientation above a surface. The only solution that covers all of the possible dimensions are SecondLight [10] and the ARToolKit [1]. LightSense [14] tracks a mobile phone with its 2D-position on a surface and a discrete 1D-distance but does not gain any orientation information. PlayAnywhere [17], BlueTable [18], iPod Party [8], reacTIVision [2], reacTable [11] and the Microsoft Surface [13] are capable of tracking mobile devices or markers with their 2D-coordinates and orientation information. Constanza's adjacency tree approach [6] is capable of tracking markers only with 2D-coordinates.

The computational needs, that are necessary to realize spatial awareness, vary from solution to solution. Generally, technologies, that involve image processing, are computationally expensive, though all of the solutions presented here run with at least 30fps. Only the hotspot tracking in BlueTable [14] and the general idea of tracking with a pattern of touch points in iPod Party [8] do not need lots of computational power.

The only tracking solution that needs greater hardware modifications is iPod Party [8]. Here a tracking carriage is needed that changes the optic and haptic of the device, that is going to be tracked. If the marker approaches are used for the tracking of devices, markers must be placed on them. But this can be done inconspicuously, for example on the bottom of the device or by the use of invisible (for example infrared) markers.

3 SUPERIMPOSED DISPLAYS

This section will present different variants of superimposed displays. We figured out three main categories, superimposing of displays can fall in: virtual superimposing, distant superimposing and direct superimposing. These categories and the most important parts of work falling into them will be presented now.

3.1 Virtual Superimposing

Virtual superimposing means, that physically only one display is available, but with this several virtual displays are created, that can then be superimposed. Firstly different variants of superimposing on one screen will be presented. Then we will have a look on virtual superimposing with Magic Lenses in Augmented Reality (AR) environments. The last topic of this section are images, that are projected through an electronically switchable diffuser, thus giving the possibility to create multiple virtual displays.

3.1.1 Virtual Superimposing on One Screen

Superimposing on one screen has been realized since the first window-based user interfaces came up. Each window in such a User Interface (UI) contains different content, and can be moved, be sent to back or brought to front. Such a window can be seen as a virtual screen. When these windows are occluding each other, this is a primitive way of superimposing. It is primitive in that way, that the single windows are usually independent from each other and do not influence each other. Special cases of this kind of superimposing displays are virtual machines or remote desktop applications, where displays that are usually meant to be displayed on a own screen are encapsulated in a window and so the contents of multiple physical displays are virtually displayed and superimposed on only one physical display.

Another way, how superimposing on one screen can be used are Magic Lenses. The first definition of a magic lens was given by Bier in 1993: "A Magic Lens filter modifies the image displayed on a region of the screen, called the viewing region, by applying a viewing filter to objects in a model." [3] That means, that filters can be applied to specific regions on the screen, that can be used for different purposes. The example in figure 1 shows a knotwork image with two different overlapping Magic Lenses applied, that stack in the region where they are overlapping. As shown later on the principle of Magic Lenses can also be ported to multiple displays or AR. There physical displays can be superimposed, both displaying the same image. Then on one display a filter is applied (This could be a visual filter, an x-ray vision or any filter that augments the underlying image with additional information).

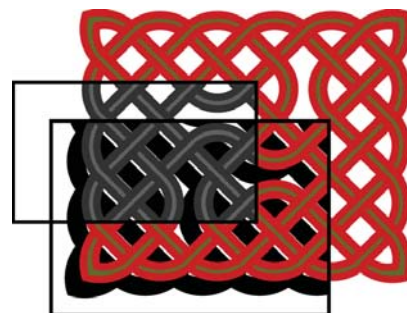


Fig. 1. An achromatic lens over a drop shadow lens over a knotwork. [3]

3.1.2 Virtual Superimposing with Magic Lenses in AR

Looser et al. [12] presented a new type of 3D tangible Magic Lens in the form of a flexible sheet, like seen in figure 2. The idea is to create a virtual Magic Lens, that can be controlled by one or two physical handles. With one handle it works like a usual Magic Lens, providing filtering of different types of data. Using two handles, the size and shape of the lens are given implicitly by the position and orientation of the handles. With two handles a variety of different dual

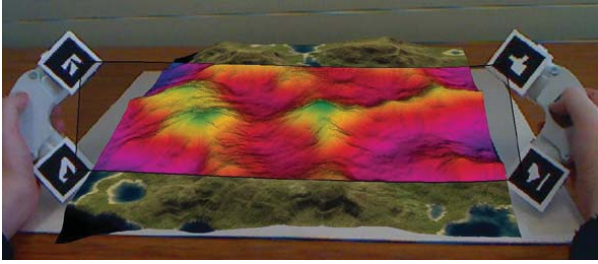


Fig. 2. Flexible Tangible AR Magic Lens. [12]

handle postures can control different user interactions. The postures, that were implemented are stretching, fanning, bending, and twisting, which can be realized by moving and rotating the two handles. The pose recognition was implemented using an artificial neural network, to whom the different poses could be trained.

Different applications were implemented to test different interaction scenarios. A geographical information system, where a 3D virtual globe is shown and the Magic Lens can be used to explore different types of data, that can be switched with a bending pose. The data layer that is shown can be selected by the amount of bending giving the user a rapid access to all available layers. Another one is an application, where the interior of physical objects can be shown and exploded for example by performing a stretching pose.

In the geographical information system, in one AR view the globe is displayed, while a virtual display – the Magic Lens – is brought over it to show additional information in different layers.

3.1.3 Projection Through an Electronically Switchable Diffuser

In 2008 Izadi et al. presented SecondLight [10]. It is a surface technology, where all features of rear projection vision systems are included, extending them for the capability of projecting through the projection screen (see figure 3). The projection surface of the SecondLight table



Fig. 3. SecondLight. [10]

is an electronically switchable liquid crystal material similar to that used in 'privacy glass', which can be switched between transparent and diffuse states. When transparent the material is clear like glass. When diffuse the material has a frosted appearance, like usual rear-projection surfaces and light, that passes through in either direction will scatter. In SecondLight this material is used as a rear-projection screen. This gives the system the ability to project on the surface, when the diffuser is in diffused state or to project through it, when it is in clear state. The diffuser is switched between clear and diffuse state at 60Hz with equal time periods for both states. This way it is possible to project on and through the diffuser at the same time. Therefore two off-the-shelf projectors were used with optical shutters attached, that are in sync with the diffuser so that all the time one shutter is clear, while the other one blocks light. This way one projector always projects on the diffuser, when it is diffused, while the other one projects through it, when it is clear.

For interaction two cameras are used, that both have IR-filters attached to only monitor IR-light. Both diffuse and FTIR IR light

sources are used in conjunction with the first camera to sense multiple fingers and other objects. This camera grabs images when the switchable diffuser is in diffuse state. The other camera only grabs images when the diffuser is in clear state, to be able to see far beyond the surface.

With this setup it is now possible to work like with a usual FTIR or Diffused Illumination (DI) multi-touch table, but extended among features of projecting through the surface. Like seen in figure 3 it is possible to project on diffuse sheets of paper, lying on top of the surface or even held behind it, making them virtual displays. These virtual displays can then be superimposed on the surface or on each other, for example to augment the images on the surface or give kind of x-ray vision through them thus turning them into a magic lens. Multi-touch interaction can also be made possible above the SecondLight table. Therefore an actively tracked transparent surface can be used. It needs an own battery-powered IR light source to realize multi-touch input on its surface, that can be sensed through the SecondLight surface.

Though SecondLight does not superimpose real physical displays, a tangible variant of virtual superimposing displays was created, that has a nearly invisible technology and a very natural feeling. And even if the basic technology is not really cheap, due to the fact, that multiple beamers, cameras and electronic shutters/diffusers are needed, it can be expanded with additional virtual (multi-touch) displays at nearly no costs.

3.2 Distant Superimposing

Distant superimposing means, that there are two or more physical displays available, that are superimposed while having a certain distance to each other. In this section we will present Shoot and Copy by Boring et al. [4] as well as Magic Lenses for AR Environments by Brown et al. [5].

3.2.1 Acquiring Information from Public Displays

Boring et al.'s Shoot & Copy [4] (see figure 4) is a system, that enables phonecam-based information transfer from public displays onto mobile phones. The project's motivation is the following. Large elec-



Fig. 4. Shoot and Copy. [4]

tronic displays become increasingly affordable and can thus be found in more and more public places. The problem is, that people hardly remember information, that they see on those public displays. Here Shoot & Copy hooks in. It is a system, that gives the user the ability to transfer information from a public display to his personal mobile phone. This is done, by just taking a photograph of the information of interest. The application evaluates, which piece of information was photographed and transfers the desired information to the user's mobile phone. At this, the system does not rely on visual codes, but it identifies the arrangement of icons on the screen and uses it to identify which piece of information was photographed by the user.

In the paper a fully functional prototype was presented, that was used to evaluate the new interaction technique. The mobile phone that was used for this purpose was a Sony Ericsson K800i with a camera resolution of 3.2 megapixels. The communication between the mobile phone and the server takes place by either GPRS or Bluetooth. The application is split into two major phases: the capture phase and the access phase. Within the capture phase the user identifies the desired information by taking a picture of it. Then a URL which points to the desired information is transferred to the users mobile phone. In the second phase, the access phase, the user can connect the mobile phone to a PC. There he can browse the URLs of the information he photographed and access them with a usual web browser. Alternatively he can also access certain information directly with the built-in web browser of the mobile phone.

The application was evaluated with 28 participants at an age between 18 and 47. The results of the study were the following: 82% said they would use Shoot & Copy if it were available. Almost all participants said that scanning and retrieving information displayed on public screens would be of interest for them. All of the users were able to use the application after a short explanation of the camera's capabilities. The users also felt comfortable while using the system. Also privacy is an interesting point to a system like the presented one: The users stated that everybody in the environment could see that they are capturing information. The regarded that nobody could see which one has been photographed.

3.2.2 Distant Superimposing in AR Environments

In section 3.1.2 virtual superimposing with the aid of AR was presented. Now Brown et al. with their Magic Lenses for Augmented Virtual Environments [5] shall be discussed. It superimposes different physical displays as Magic Lenses in an AR space. The display setup presented within this work describes a wall sized room display, a workbench display (the so-called World in Miniature or WIM) and two different kinds of Magic Lenses: a handheld Magic Lens and an embedded Magic Lens. The handheld device is a mirror-like, movable prop with a 5" inset surface serving as the presentation window. Spatial awareness is realized with an Ascension tracker, that provides position and orientation information of the device relative to the workbench reference. It is equipped with some controls to make interactions beyond moving and rotating possible (for example context switches or altering zoom levels). The second lens interface decouples the physical interface components used for data inspection and Point of interest (POI) selection. A large 14" inspection window, is attached directly to the workbench and remains stationary throughout operation. The focus window is disjoint from and situated to the right of the WIM. POIs can then be selected from the WIM interface by using a cordless, physical prop called a selection icon. ([5], p. 6)

Two testbed applications are presented: Aztec Explorer and Anatomy Explorer, that simulate probable applications in urban planning and medical training. Aztec Explorer lets users explore an ancient Aztec city. The Scape room display, how the wall sized room display is called, presents the life-sized, highly detailed city while the workbench houses a corresponding low-detail WIM interface. The Magic Lens devices serve as a bridge between these dual scales and perspectives. Anatomy Explorer allows the inspection of the interior of the human torso and abdomen. The workbench displays a life-sized view, while the room display presents a corresponding molecular fly-through view. This application provides exciting augmentation capabilities, as users can overlay the data sets onto an actual human patient, simulating a composite 3D augmented view into that persons body. The handheld Magic Lens can be used by the user to peer inside the person's physical body. Organs can be inspected closely with the aid of the zoom feature, while a clipping plane can be used to selectively cut away occluding organs. The filtering mode can be invoked to turn various organ systems on or off. These two applications can be seen in figure 5.

The setup was evaluated with five students, who had to find three markers, that were hidden in the Aztec Explorer scene and though not occluded from the students field of view, were so small, that a Magic

Lens was needed to find them. The markers had a letter imprinted on them, the students had to identify. The result was that the students preferred searching with the handheld Magic Lens because it was faster and more intuitive to use, but preferred the embedded Magic Lens for reading the imprint, because it did not suffer so much from hand jitter.

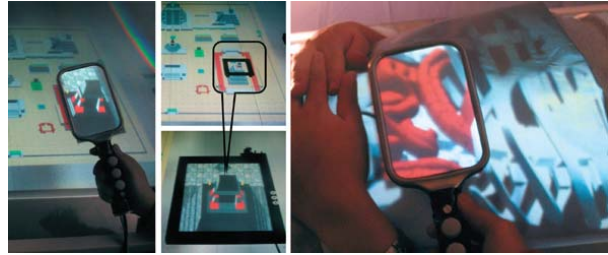


Fig. 5. Left: exploring temples in an ancient aztec city. Right: exploring the anatomy of a human torso and abdomen. [5]

3.3 Direct Superimposing

Direct superimposing means, that two or more displays are superimposed directly on top of each other, thus touching each other or having a very low distance to each other (not more than one or two centimeters). Three different ways of direct superimposing will be discussed now. Firstly a look on the applications that came with LightSense [14] will be taken. These superimpose a mobile phone on a static map. Following that, two different works regarding the superimposing of displays to enhance the local display resolution will be presented. The last thing is superimposing of displays to enhance the overall display resolution, what can be realized by superimposing the displays of several projectors in a multi-projector environment.

3.3.1 Direct Superimposing to Augment Static Maps

The LightSense [14] tracking technique was already introduced in section 2.1. Now we will have a look on the applications, that were also presented in [14] (see figure6).



Fig. 6. Left: a wall-mounted map with embedded light sensors provides hotspot tracking. Right: physical map of the Stockholm subway is augmented with digital content. [14]

The first application, that was presented, augmented a physical map of the Stockholm subway with digital content in a focus and context fashion. The system is able to identify the closest station to the device. Then it can provide the user with a local map for the station's surroundings. It also gives him the ability to zoom in and out by varying the distance of the phone to the physical map. The potential to use this application to aid disabled people, like vision impaired users, is also given. Therefore the mobile phone can be used as a scanning device, that aids users in following a certain subway line, by giving vibrotactile feedback. When the user loses contact to the a subway line, he is exploring, the phone starts vibrating. The names of stations, the phone moves over, can be played back via the ear phones and it would also be possible to read out colors for color blind users. Non-disabled users could use this application as a personal magnifying glass. The application uses techniques according to Yee's Peephole Displays [20] or

the work of Fitzmaurice [7]. Olwal says: “The device acts as a viewport into a large image, for example and as the user moves the display around, different portions of the larger image will be seen.” ([14], p. 121) LightSense units could be used as ubiquitous guides in such a way, that traditional media, like posters, could be equipped with LDR sensors and so giving them the ability to serve additional information, when a mobile phone is superimposed to them.

3.3.2 Superimposing Displays to Enhance Local Resolution

Two displays with different dimensions and different spatial resolutions can be superimposed to enhance the local resolution of the display with lower spatial resolution. Two different approaches to this were shown in Olwal’s Spatially Aware Handhelds for High-Precision Tangible Interaction with Large Displays [15] and iPod Party [8]. Both of them took a mobile device with touch screen (Sony Ericsson G900i and Apple iPod touch) and lay it on a multi-touch table to use the mobile devices display to enhance the local spatial display resolution of the table. To realize spatial awareness Olwal used the LightSense [14] technique discussed in section 2.1, while in iPod Party the tracking by generating a pattern of touch points discussed in section 2.3 was used.

Olwal’s hardware setup contains a 20" touch-sensitive table with a display resolution of 800×600 pixels and a camera below the projection surface for the spatial awareness of the mobile device. The Sony Ericsson G900i has a display resolution of 240×320 and is single-touch-sensitive. In this setup the visual and input resolution per unit is about 11 times higher on the mobile phone than on the rear projected touch-screen surface. Different interaction techniques, all for zooming, were presented: four different interaction techniques on the mobile phone and two on the table. The mobile device is only controlled by the stylus, while the table is controlled with fingers. The zoom techniques that were invented for the mobile phone are MobileRub, MobileGesture, MobileDrag and MobileButtons. Items can always be selected by clicking them with the stylus. MobileRub works with small diagonal rubbing gestures over the target, while maintaining contact with the screen. The slope of the diagonal rubbing causes the action to be taken: rubbing along the diagonal with a positive slope zooms in incrementally for each stroke, while rubbing along the diagonal with negative slope resets the zoom level. Along MobileGesture, zooming is controlled with circular gestures. Counterclockwise circles zoom up to the extent of the circle, while clockwise circles reset the zoom level. In MobileDrag, zooming is controlled by dragging gestures. The user zooms in on a specific position by touching it and dragging upwards or downwards, to zoom in or out. The change on the zoom factor is controlled by the distance the user has dragged. MobileButtons does not use the touch capabilities of the mobile phone, but works with the phone buttons. The up- and down-buttons alter the zoom level, while items can be selected with the center button. On the table Zoom-Pointing and Rub-Pointing were implemented. The zoom mode can be entered by pressing a zoom button. After entering the zoom mode, the user can drag out a rectangle around the target. The System zooms to the area specified by the rectangle. This can be repeated until the target is large enough to be clearly selected by direct touch. Rub-Pointing is similar to MobileRub, but instead of on the phone’s display rubbing takes place with the finger on the table’s surface. The interaction techniques were evaluated with twelve students. The result was, that the mobile techniques were preferred before the on-table techniques. Among the mobile techniques MobileGesture was ranked highest, followed by MobileDrag and MobileRub, with MobileButtons ranked lowest.

The aim of iPodParty [8] was, to create an application, that can be used to analyze the benefits of the combination of a multi-touch tabletop display with one or more multi-touch handheld displays. This was done by superimposing them to enhance the local input and display resolution on the table. A photo of the application can be seen in figure 7. The hardware setup of the application contains a multi-touch tabletop display with a display resolution of 1024×768 and a diagonal of approximately 50". The multi-touch input is realized with FTIR [9] with a camera with a resolution of 640×480 pixels. Additionally one or more iPod touches can be used, that have a display diagonal of

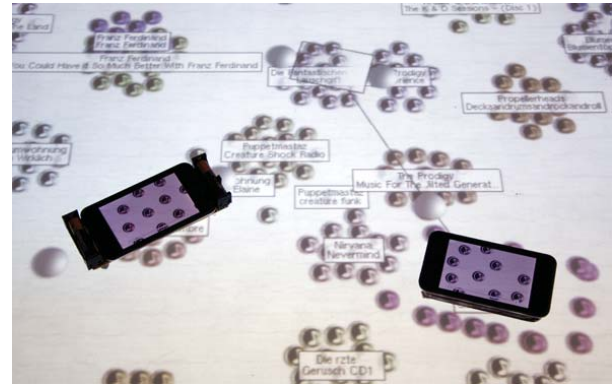


Fig. 7. iPodParty with two iPod touches. [8]

8.9" at a resolution of 480×320 pixels and are capable of recognizing multi-touch input. Four different interaction modes were introduced, two with the handheld directly superimposed to the multi-touch table and two, where the handheld is held in the users hands. Firstly the handheld as magnifying glass is presented. That is just to place the handheld on the table and use it as a magnifying glass for the underlying information. With this method the higher resolution of the handheld can be used to gain more detail in specific sections of the tabletop display. The magnification level can be altered with a two finger zoom gesture on the handheld’s display. The second mode is the handheld with magnifier arm. With this mode, it is possible to pull a magnifier arm out of the handheld. This can be visualized with a rectangle that is connected to the handheld via a thin line. This rectangle represents the area of magnification, whose content is shown on the handheld’s display. With this abstraction the user can place the area of magnification where he wants, even if he can not reach it with the handheld. The other two modes correspond to the presented modes, but with the handheld held in the users hands. Instead of itself, an anchor point on the table, represents its position. The iPodParty application presented in that these is an audio browsing application, where the songs of an audio library are represented by small icons, that are ordered by similarity on the multi-touch table and can be browsed using an iPod touch. There was no user evaluation in iPodParty, so no conclusion can be given on how good the interaction concepts work and which ones are the best.

3.3.3 Superimposing Displays to Enhance Overall Resolution

Multiple displays can also be superimposed to enhance the overall display resolution. This is done in PixelFlex [19] and PixelFlex2 [16] by using multiple projectors whose areas of projection slightly overlap and who are fast reconfigurable.

The configuration of the PixelFlex system contains the following hardware components: a control PC, a camera and eight Proxima 6850 LCD projectors driven by multiple pipes of an SGI Infinite Reality system. The projectors have a resolution of 1024×768 and output 1500 ANSI lumens. In front of every projector, a mirror is mounted on a pan-tilt unit (PTU), what gives the ability to electronically adjust the position of the projectors projected image. The projectors with the mirror-PTUs are mounted to the ceiling in an array of two rows of four. Several configurations are possible, but in the typical wide-area configuration the whole display area is approximately 12 feet by 5 feet, with an average spatial resolution of 25 DPI and 15% overlap. The projector layout can be changed from the control PC by adjusting the mirror-PTUs. It is possible to save and load any configuration made. After each layout change a calibration must be made. The calibration is done in two main steps: geometric registration and afterwards photometric calibration. The geometric registration is described the following by Yang et al.: “The goal of the geometric registration procedure is to create a mapping between each projector’s image coordinates and the display’s global coordinates. There are three main steps

in the geometric registration process: (1) camera to display surface registration, (2) projector registration via structured light, and (3) post processing of the registration data for the appropriate rendering algorithm” ([19], p. 3) After the geometric registration, the multi-projector environment can be used as one large display, but in the regions where projected images are overlapping, the light intensity is higher than in regions where only one projection is located. To correct this, photometric calibration is to be made (see figure 8). Yang et al. describe it

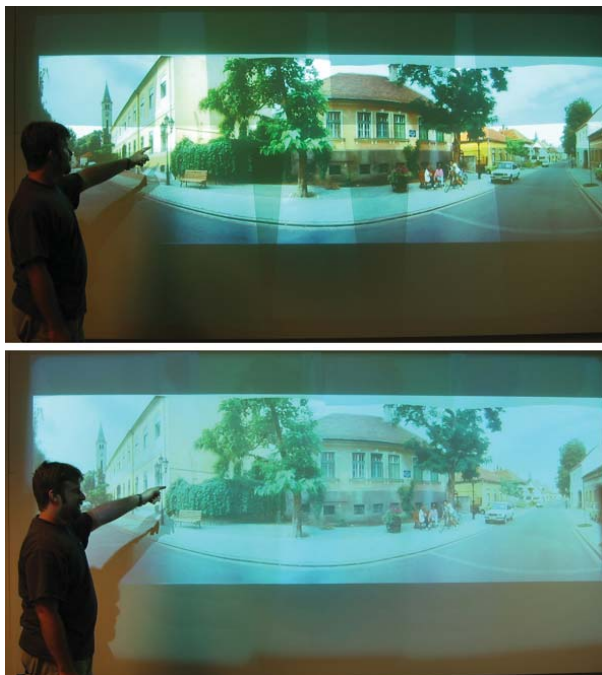


Fig. 8. The same image before (top) and after (bottom) photometric correction. [16]

like this: “There are two major tasks involved with photometric calibration. The first task is the measurement of each projector’s intensity response. This data is used to create a color lookup table that linearizes each projector’s intensity response. The second task is the determination of the display overlap regions. This data is used to compute an alpha (blending) mask used in the rendering process to attenuate the light contribution of individual projectors in these overlapping regions so as to produce a photometrically seamless display.” ([19], p. 4). Now, that the calibration process is finished and a large-scale, high-resolution display is available, a rendering application is needed. For PixelFlex two different rendering applications were implemented: an X Window Desktop, realized using a modified version of the VNC remote desktop software, that supports most X Window Applications. Unfortunately it does not support OpenGL programs, so an additional OpenGL 3D Viewer application was created. With these rendering applications and a wide area configuration of PixelFlex a resolution of approximately 3500×1300 Pixels can be used.

Though PixelFlex seems to work quite well, it has still some issues, where PixelFlex2 hooks in and brings some improvements. It was migrated from the SGI graphics system to a Linux PC-cluster for higher performance and scalability. It has an improved geometric registration process, because it uses a checkerboard corner-finder instead of a Gaussian centroid-finder. Also the photometric calibration was improved. In PixelFlex the intensity response needed to be measured manually with a Photometer, now it is calculated automatically with HDR-images that are computed from the images of the calibration camera. The attenuation solution realized in PixelFlex2 is completely image-based, so no need to know the exact locations of the overlapping regions is present. Another new feature is a black offset compensation mask, that guarantees a constant black level over the

whole display surface.

So with PixelFlex and PixelFlex2, solutions were presented, that can increase the overall display resolution from 1024×768 pixels of one projector to about 3500×1300 Pixels of eight projectors by superimposing them. This is an increasing of a factor of about 6.

4 CONCLUSION

In this paper an overview on the current state of research in the field of superimposed displays was given.

Firstly we gave an overview over tracking technologies, that are needed to realize spatial awareness, what is important to many applications, that work with superimposed displays. The technologies that were presented were tracking with the photolight of a mobile phone, tracking with a camera mounted in front of the superimposed displays, tracking by generating a pattern of touch points and tracking with markers. After presenting these technologies, a short comparison of them was given, regarding aspects like robustness to external influences, tracking dimensions, computation needs or needed hardware modifications.

Afterwards, there was the main part, where actual papers, regarding topics for superimposed displays were presented. This part was divided into three subparts: virtual superimposing, distant superimposing and direct superimposing. Among virtual superimposing, firstly virtual superimposing on one screen was discussed, regarding different possibilities like windows, virtual machines and remote desktops, as well as the Magic Lens principle presented by Bier in [3], that was extended and ported to other domains later on. The next section covered one of these domains: Magic Lenses in AR. A flexible tangible AR Magic Lens, presented by Looser [12] was presented. Afterwards we came to SecondLight, a technology by Izadi et al. [10], that makes it possible to project through an electronically switchable diffuser and thus making it possible to virtually superimpose displays. Then the next subpart was to come: distant superimposing. We discussed Boring et al.’s Shoot & Copy, that can be used to receive information from public displays with the aid of the camera of a mobile phone. Brown et al.’s [5] Magic Lenses, that can be used in a mixed AR space, were presented, as well as the applications Aztec Explorer and Anatomy Explorer. The last subpart, direct superimposing, started with static maps, that are directly superimposed with a mobile phone, to augment them with additional information. After that Spatially Aware Handhelds for High-Precision Tangible Interaction with Large Displays [15] and iPodParty [8] were presented, that both deal with the possibilities to superimpose small devices on tabletop device to locally enhance their display and input resolution. The last topic that was presented is superimposing of displays to enhance the overall display resolution, what was done in PixelFlex and PixelFlex2 by overlapping the projection areas of multiple projectors to achieve a high-resolution projection.

REFERENCES

- [1] ARToolworks, Inc. ARToolKit. <http://www.hitl.washington.edu/artoolkit/>, 2009. visited 15.06.2009.
- [2] R. Bencina, M. Kaltentbrunner, and S. Jorda. Improved topological fiducial tracking in the reactivation system. In *CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, Washington, DC, USA, 2005. IEEE Computer Society.
- [3] E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose. Tool-glass and magic lenses: the see-through interface. In *SIGGRAPH '93: Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 73–80, New York, NY, USA, 1993. ACM.
- [4] S. Boring, M. Altendorfer, G. Broll, O. Hilliges, and A. Butz. Shoot & copy: phonecam-based information transfer from public displays onto mobile phones. In *Mobility '07: Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology*, pages 24–31, New York, NY, USA, 2007. ACM.
- [5] L. D. Brown and H. Hua. Magic lenses for augmented virtual environments. *IEEE Computer Graphics and Applications*, 26(4):64–73, 2006.

- [6] E. Costanza and J. Robinson. A region adjacency tree approach to the detection and design of fiducials. In *Video, Vision and Graphics*, York, United Kingdom, 2003.
- [7] G. W. Fitzmaurice. Situated information spaces and spatially aware palm-top computers. *Commun. ACM*, 36(7):39–49, 1993.
- [8] S. Gebhardt. iPod Party: Designing an application to explore the possibilities of the combination of a tabletop display with one or more handheld displays. Project Theses, LMU, Munich, 2009.
- [9] J. Y. Han. Low-cost multi-touch sensing through frustrated total internal reflection. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 115–118, New York, NY, USA, 2005. ACM.
- [10] S. Izadi, S. Hodges, S. Taylor, D. Rosenfeld, N. Villar, A. Butler, and J. Westhues. Going beyond the display: A surface technology with an electronically switchable diffuser. In *UIST '08: Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 269–278, New York, NY, USA, 2008. ACM.
- [11] M. Kaltenbrunner, S. Jorda, G. Geiger, and M. Alonso. The reacTable*: A collaborative musical instrument. In *Enabling Technologies: Infrastructure for Collaborative Enterprises, 2006. WETICE '06. 15th IEEE International Workshops on*, pages 406–411, 2006.
- [12] J. Looser, R. Grasset, and M. Billinghurst. A 3d flexible and tangible magic lens in augmented reality. In *ISMAR '07: Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 1–4, Washington, DC, USA, 2007. IEEE Computer Society.
- [13] Microsoft. Microsoft Surface™. <http://www.microsoft.com/surface>, 2009. visited 27.05.2009.
- [14] A. Olwal. LightSense: enabling spatially aware handheld interaction devices. In *ISMAR '06: Proceedings of the 2006 Fifth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'06)*, pages 119–122, Washington, DC, USA, 2006. IEEE Computer Society.
- [15] A. Olwal and S. Feiner. Spatially aware handhelds for high-precision tangible interaction with large displays. In *TEI '09: Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*, pages 181–188, New York, NY, USA, 2009. ACM.
- [16] A. Raij, G. Gill, A. Majumder, H. Towles, and H. Fuchs. Pixelflex2: A comprehensive, automatic, casually-aligned multi-projector display. In *In Proc. IEEE International Workshop on Projector-Camera Systems*, 2003.
- [17] A. D. Wilson. PlayAnywhere: a compact interactive tabletop projection-vision system. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 83–92, New York, NY, USA, 2005. ACM.
- [18] A. D. Wilson and R. Sarin. BlueTable: connecting wireless mobile devices on interactive surfaces using vision-based handshaking. In *GI '07: Proceedings of Graphics Interface 2007*, pages 119–125, New York, NY, USA, 2007. ACM.
- [19] R. Yang, D. Gotz, J. Hensley, H. Towles, and M. S. Brown. Pixelflex: a reconfigurable multi-projector display system. In *VIS '01: Proceedings of the conference on Visualization '01*, pages 167–174, Washington, DC, USA, 2001. IEEE Computer Society.
- [20] K.-P. Yee. Peephole displays: pen interaction on spatially aware handheld computers. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1–8, New York, NY, USA, 2003. ACM Press.

Applications for interactive public displays

Miriam Kranz

Abstract— Until now public displays were simply broadcasting centrally produced media content to a large audience. With the improvement in display and sensor technology it is possible to create large, interactive displays for public spaces. Therefore the concept of applications on public displays must be over thought and adapted to this new situation. Many projects have already dealt with this theme which resulted in a multitude of new applications for such interactive public displays. In this paper an overview of these applications and some study results will be given and they will be analyzed and compared in order to detect which usability features those applications need to be noticed, accepted and used by the public. The resulting characteristics could serve as guidelines, ease the use and development of further applications or initiate additional research.

Index Terms—interaction, applications, public displays, application features, interactive surface, usability

1 INTRODUCTION

1.1 Motivation

Nowadays public displays are exclusively used for the distribution of centrally produced media content that is directed to a large audience, as for example information screens at airports or in metro stations [5].

At the same time the increased development of mobile devices enables customers to perform more and more of their private activities in public settings. Talking on the phone or surfing the internet and sending mails are no longer activities that require a terrestrial connection but can be performed almost anywhere at anytime. Eriksson et al. declare that in public spaces the information exchange is more an information push then information dialog and instead of social interaction there is a development toward a more personal sphere [10].

Nevertheless, when it comes to gather information about upcoming events or share information with other members of their community, people often rely on non-electronic public surfaces such as posters, bulletin boards or fliers. However, there are several restrictions bound to those non-electronic public displays like the impossibility to remotely modify content or the limitation of size [2].

With the improvement in display and sensor technology it is possible to create large, interactive displays for public spaces that could solve lot of those problems, make people become producers and users, not only consumers and rebalance the information circulation in public spaces.

1.2 Definitions

In the term *applications for interactive public displays* there are lots of concepts that need to be defined to get a better understanding of the theme.

First, what does **interaction** mean in such a context? When analyzing the interaction possibilities an application provides, we will distinguish between different interaction ways, degrees and tools.

Here, two main ways of interaction can be discerned: implicit and explicit interaction. On the one hand there is the implicit interaction where the system reacts on certain signals like the body orientation of a passer-by or the proximity of a person to the display. This gives the user the impression the system interacts with him, not him interacting with the system. On the other hand there is the explicit way where the user interacts with the display through explicit actions such as touching the screen, hand gestures or using the mobile phone to send a message to the system. Often the implicit way of interaction is used to get the attention of a passer-by and move him to use the display in an explicit way [24].

An application can offer different degrees of interaction. Some applications are limited to the consultation of information; the user can browse data, view or sometimes even sort it [19]. Still there is no possibility for him to modify, update, upload or download the information. Some other applications offer there users the additional possibility to modify or update data. Here users can for example annotate content, leave comments or modify the screen in an interactive game [4, 7]. An even higher degree of interaction is given when the application offers the user the additional possibility to upload and/or download data from/to mobile devices [15, 5].

An other interactivity point from which many applications differ are the possible interaction tools. These tools could be PDAs, mobile phones, handheld controllers, the users' hands, USB-sticks, mice, keyboards and many other. Whereas some applications only enable the interaction with the system through one tool as for example mobile phones [20, 12] others offer the user the possibility to use a multitude of different tools [15].

What is an **interactive public display**? In [15] it is defined as "shared surface used for different tasks by an open-ended community of users". The aspects that discern the interactive *public* display from a *regular* interactive display are the location and the users. While most applications on regular displays are designed for a specific target audience, the users from public displays form a heterogeneous mass. For this reason an application on such a display has to deal with popular and well known themes to appeal to its large audience. Public displays are situated in city streets, shopping centers or public buildings like hospitals, hotel lobbies, university hallways and other public places.

Nevertheless, the technologies of public and non-public interactive displays remain the same. An interactive public display could for example be a plasma display with a touch overlay like in [6] or a surface on which an application interface is projected and where interaction is tracked using a camera based system [4]. This paper will concentrate on the applications that are developed for such interactive public displays.

1.3 Paper Content

The following paragraphs aim at giving an overview of the different applications displayed on such interactive public surfaces. Besides some general properties that contribute to the acceptance of an application by the audience will be elaborated.

Therefore some applications from different domains will be described, analyzed and important features will be extracted. Finally a recapitulation and explanation of the different usability properties will sum up the findings from the analysis of the applications.

2 APPLICATION ANALYSIS

In this section seven applications from different domains will be presented and analyzed. Therefore a quick overview of the goal and main features of each application will be given. Special attention will be paid to how interactive the application is by examining the interaction

-
- Miriam Kranz is studying Media Informatics at the University of Munich, Germany, E-mail: miriam.kranz@campus.lmu.de
 - This research paper was written for the Media Informatics Advanced Seminar on Interactive Surfaces, 2009

ways, degrees and tools. A short presentation of some results of the studies linked to each application will underline the achievements and encountered problems. This information contributes to the identification of some general features such an application should have. Those are extracted and explained afterwards. The applications are sorted by complexity and interaction possibilities.

2.1 CityWall: explore images

CityWall [19] is an application designed for large multi-touch displays that are installed in city centers. Passers-by can explore and play with pictures. The users can select pictures from Flickr and then resize, rotate and move them through simple hand gestures. The application has a zoomable timeline on which the images can be organized chronologically.

The interaction possibilities are restricted: the only interaction tools are the users' hands and there is no possibility to upload or download content from/to the users' mobile devices. Even the possibilities to modify the content are limited. There is no implicit interaction with the display.

In a study with CityWall [19] the display was installed in the center of Helsinki and observed for eight days. The results showed that the application was used only 8.8 percent of the total uptime, but seeing other people interact with the display served as attractor for more people to come. In figure 1 a) people seek shelter from the rain next to the application without noticing it. Figure 1 b) shows how a boy notices the CityWall and immediately draws the attention of the surrounding bystanders to the application.



Fig. 1. Passers-by seeking shelter from the rain a) and interacting with the CityWall behind them b) [19]

The users almost never needed to consult the instructions printed next to the screen, watching other people interact with the system was sufficient to understand how the application works. Sometimes conflict situations occurred like overlapping and blowing up the picture of another user or using the timeline irrespective of what the other users were doing. The concerned users often adopted an unapprising posture and immediately stopped to use the display.

Considering the information above, some important features can be brought to light. First the need to **attract attention**. In modern cities the visual senses are overloaded with information, in order to not remain unnoticed the applications need to attract the attention of the passers-by. This can be seen in the described study, where people don't pay attention to the display when nobody is using it. Here people using the display piqued the curiosity of other bystanders.

Another deduction that can be drawn is the importance of **conflict management**. When many individuals are interacting on the same surface, conflicts are inevitable as can be seen in the study described above. These situations could be many users simultaneously trying to access the same resource or one individual trying to close the resource of another one. Therefore the application must contain methods to resolve such conflict situations.

2.2 GroupCast: create conversation opportunities

GroupCast [17] is used to enrich casual interaction. The application runs in the background and aims at creating conversation opportunities, showing common or exclusive interests of persons standing next to the screen. A possible scenario could be two persons passing in front of the screen. The 'Wine of the day' website pops up leading to a spontaneous discussion about the different steps in the production of wine [17]. At the beginning, the developers worked on a form people should fill out in order to gather information about their interests

but it soon turned out that no one would take the time to do this. In parallel, they developed an application called UniCast [16] where user profiles were generated and decided to use the interests indicated in those profiles as potential conversation themes.

There is no direct interaction with the display, only implicit interaction. Possible interaction tools are infrared badges or radio frequency identification (RFID) tags. In some way the user also interacts with the display when changing his UniCast profile but the interaction possibilities still remain very restrained.

McCarty points out that one of the main problems of such applications is how to acquire profile information about the people passing by the display. This is also due to privacy concerns and social embarrassment. Another problem is how to attach RFID-tags or infrared badges to passers-by.

This analysis leads us to two other important features that should be considered when developing an application. First the need of **privacy control**, an application must discern private data from public data. Nobody would use an application on a public screen when he runs the risks of giving away private information in public. Another point is **social embarrassment**. People often feel embarrassed when performing actions in public, where everybody recognizes them. In this setting for example, if two people pass the display and it indicates an interest of one of the passers-by that the other one totally disapproves, this could lead to social embarrassment.

Another application that deals with creating conversation opportunities is the Interactive Wall Map [17]. It consists of a wall map of the world with integrated touch screens and LED topped button switches. Like GroupCast, it runs in background and uses the UniCast profile framework to index content geographically on one of the integrated touch displays. But it also allows explicit interaction by pressing LED buttons which brings up content related to its position. This additional possibility to explicitly interact with the system gives the users the opportunity to take a less passive role. Also the fact that the displayed content is only about geographical regions reduces the social embarrassment and privacy concerns.

2.3 MobiLenin: a voting system

Several applications based on a voting system for interactive public displays have been designed. For example UniVote [8], an application developed to enhance user interaction with the campus and outside world, by informing the students about campus- and world-wide events or news; or MAGICBoard [22] which gathers opinions of bystanders on amusing topics. In this section the focus will be set on the voting application MobiLenin.

MobiLenin [20] is a voting system that allows a group of people to interact simultaneously with a multi-track music video displayed on a large public screen and it aims at enticing social interaction between people. The users can vote for a clip by using a client application installed on cell phones. The system also integrates a lottery mechanism. The large public display indicates the start and end of a voting period as well as the voting results and it displays the voted track (see figure 2).



Fig. 2. A screenshot of the MobiLenin public display showing the voted track and the voting results [20]

The users can only interact via cell phones. In addition the application on the mobile device only gives them restricted possibilities, they can solely choose between a few tracks. There is no possibility to upload, download, update or modify content.

The system was tested in a restaurant where mobile phones with the client application were distributed among the customers and instructions were given on how to start the application. Through the lottery system the user was given the chance to win a pizza or a beer. The study showed that application was found easy to use, popup notes on the client UI were found very useful and the lottery system motivated the users and created additional suspense. Another important finding is that the application stimulated inter-personal social interaction in co-located groups. The testers stated that they felt being a part of a group and for many of them this was even the most important point and motivation to use the application. But the users also mentioned that the system got boring after some time.

The results allow the introduction of a new feature, the **increase of a sense of community**. Because of multiple people interacting with the same application, the interaction with co-located players is facilitated. Often the pursuit of a common goal generates a sense of community, the users feel like belonging to a group. In this example the voting mechanism strengthens their sense of community. The results also underline the importance of feedback and motivation. If the users have the possibility to win something they get more motivated to participate and this also serves as attention attractor. One problem is that the application must be installed on the cell phone before; this in the contrary could reduce the motivation of users to interact with the system. Also the limited interaction and voting possibilities made the users rapidly lose interest.

2.4 Polar Defense: an interactive game

There are already a lot of interactive games developed for public displays as for example the Red Nose Game [18], Manhattan Story Mashup [23], Flashlight Jigsaw [4] or Polar Defense [12]. In this section only the last two applications will be analyzed, concentrating on Polar Defense.

Polar Defense [12] is an interactive game where users have to place towers on a screen in order to prevent as many attackers as possible from walking across the screen. The users interact with the screen via SMS messages, sending the coordinates of the towers that need to be positioned to the system. The application surface is divided in two. The information display shows an animation (*located on the right side of figure 3*) and information about the game (*located on the left side of figure 3*) to attract the audience, engage bystanders at the glance state and provide all the information needed to understand how the game works. The game display concentrates on the gameplay and



Fig. 3. The Polar Defense information display [12]

indicates the current player and the player queue via their partially hidden phone numbers, the high score and the actual game. The only possible interaction tool is the cell phone.

Here the interaction possibilities are also quite low, the user can not really upload nor download content. There is only explicit interaction and the only tool that can be used to communicate with the system is the mobile phone.

For a study the large display game was deployed at the main campus of the University of British Columbia for four days. During this period logs of user interaction data were retained and the public space around the setting was observed. Findings of that study were that the animation drew the students attention to the game. The instructions were easily understood except for the number of towers that could be placed on the field, but frequently players explained the game to others so that mistakes did not occur often (20 messages out of 203 contained errors but could be corrected by the game logic).

Here it gets clear that it is important that **instructions on how to interact with the system need to be clearly visible** to the user. If not, the player would be left in the dark without knowing how to use the application. Especially when the application principle is not as simple as the CityWall's. In case of Polar Defense the information display provides all help needed to understand how to interact with the system.

Another interesting point is the possible reduction of **social embarrassment** by employing interaction via mobile phones. The partly hidden phone number allows the user to stay anonymous if wanted. The animation and the information display attract the attention of passers-by which are new potential players. It also contributes to an easy understanding of the application. In addition to the animation the game principle that is very similar to *Tower Defense*, a well-known game genre, could also lead to a faster understanding of the gameplay.

An application where the social embarrassment problem is not resolved is Flashlight Jigsaw [4], a multiplayer puzzle game played on a large public display using wireless handheld controllers. Here the players can clearly be identified because they are holding the controller and that way are completely exposed. Some players stated that they felt nervous and intimidated when performing in front of a large audience, especially when they were performing badly. Some persons did not play at all and stated social embarrassment as reason because they "didn't want to seem dumb" [4]. Once again the importance of avoiding social embarrassment is demonstrated.

2.5 Dynamo: support collaborative work

There are multiple applications that deal with collaborative work, for example IM Here[14] an application for shared instant messaging which facilitates the communication between members of a work-group. The Notification collage [13] is an electronic bulletin board where colleagues can post different media elements like videos, sticky notes or web page thumbnails. In this section we will concentrate on the Dynamo system [15], which stands out due to its diversity of features and interaction possibilities.

Dynamo [15] supports collaborative work by allowing multiple users to simultaneously interact on the same surface and share, display or exchange media. The surface is composed of one or more displays that can be arranged horizontally or vertically. The application is designed for collaborations that take place in unfamiliar public places; a meeting at an airport or in a hotel could be a possible scenario. The users interact with the surface through interaction points where data can be uploaded or downloaded from/to USB-stick, laptop, desktop and updated or modified with mouse and keyboard. Each interaction point has its own *telepointer* and *personal palette*(see figure 4)), both marked with the same color. Small icons on the personal palette represent *media sources* and *media sinks* from which the user connected to this interaction point can upload data (sources) or on which data can be downloaded and stored (sinks). A *public palette* provides access to key features of the system such as browsing the web or creating notes. Another design element are the *carve regions*, which allow users to engrave a part of the display. The user that created the region can allow other users to dispose of that carved surface. *Media parcels* can be left on the screen (*the window entitled 'Parcel for Mia'* in figure 4) or putt into packages (*the package entitled 'Parcel for Bob'* in figure 4).The user also has the possibility to seal a parcel, so that the content



Fig. 4. Dynamo interface with a selection of media, two carve regions, three personal palettes and a public palette [15]

can not be viewed on the public surface and must be downloaded and viewed on a personal device by the person it is addressed to.

Dynamo offers many interaction possibilities, users can browse, upload, download and modify content. The possibility is given to use a multitude interaction tools, like keyboard, mouse, laptop, desktop and other tools that can be connected via USB.

In [15] two studies with the Dynamo system and their results are described. The first study took place at a hotel conference center where people were attending a workshop. It was based on an earlier version of the system without access control and the users could carry out any interaction. The study showed that carving the regions needed explicit instructions, as did dragging media onto the surface. Also some conflicts occurred when multiple users interacted simultaneously and the freedom for any user to manipulate any region on the surface gave cause for concern. These results led the developers of Dynamo to overwork the system and refine carving and access control. A second study, where the attendees were divided into small groups and given tasks to resolve, showed that the efforts of overworking the application bear fruits. Overlapping was less frequent and when they occurred they could be solved way more easily.

Regarding this information, it gets clear once again how important an **easy understanding of the application** and **conflict management** is. By improving the interaction techniques no explicit instructions are needed anymore to understand the functions of the application. This improvement saves time and leads to a more efficient collaboration. Through the access control lots of conflicts can be avoided. The application also considers **privacy control** by giving the user the possibility to seal his media parcels and by not allowing users to access the connected private devices of other users.

2.6 Plasma Poster: an interactive community board

The PlasmaPosters [6] are interactive public posterboards that facilitate informal content sharing between a group of people and consist of plasma displays with interactive overlays enabling direct touch interaction. Together, they form a network of interactive community boards. Content can be posted via a web interface and by email. Users can browse posted web content or images, read text files or display videos, by directly interacting with touch display. Additionally there is the possibility to annotate content via PDAs. Several other applications based on the Plasma Poster system, but with different interfaces,

were developed like YeTi [25] or the eyeCanvas [7]. The following passage concentrates on the Plasma Poster application as it described in [6].

The Plasma Poster interface is divided in several parts (see figure 5), the *content region* in the center of the screen displays content (images, text,...) an author has uploaded. Below this region a picture of the author and his comments about the posting are displayed. The user can send the author a note that he scribbles on the surface or browse other of his postings. Below the authors information region are the overview thumbnails that indicate items that have recently been used, are currently viewed or that are about to be viewed when the interaction with the display has ceased.

The application has a high degree of interaction; different types of media content can be uploaded, modified and annotated. But most of the work with the content does not occur on the Plasma Poster itself but via email. On the Plasma Poster users are mainly browsing through the content and sometimes leaving notes. Interaction tools are PDAs, the web or the hands and there only is explicit interaction.



Fig. 5. The Plasma Poster interface [5]

In [6] three Plasma Poster are placed in a lab at FX-PAL[www.fxpal.com/], one in the corridor, another in the foyer and the last one in the kitchen and eating area. Data is collected during

20 months of continuous use, furnishing information about the interaction and reading practices of the users and factors for the success of the application. For example the introduction of authors pictures and the thumbnail overview led to an increased use of the display. Also the application was easy to use because the ways of interaction fitted with existing practices and concepts like navigating through a forum or leave comments under a blog entry. But the Plasma Posters needed continuous service and complains occurred when the services where interrupted. Also most of the users did only read the content but did not add some, saying they do not know what to post, this could also be an indicator for social embarrassment.

Those results also point out how important it is for an application to be **easy to use** and that the **increase of a sense of community** contributes to the success of an application. Because of its interaction techniques being similar to other well-known practices hardly any instructions were needed and because of its similarity to blogs and forums the users felt like being a part of a community. This also points out a weakness of the application: it does not respond to a target group with little or no knowledge of structures like blogs or forums.

2.7 Context-aware Hospital Information System: support medical work

In [11] Favela et al. discuss the design of an application for context-aware public displays in a mobile hospital information system. The context-aware hospital information system (CHIS) forms the base. PDAs or handheld table computers allow the medical personnel to access electronic patient records, to send messages, locate patients, other staff members and resources. The system was evaluated by 28 hospitals with very positive feedback (91 percent of the participants would use the system), still the handheld table computers have significant limitation in visualizing medical images, which could be resolved by the introduction of a public display that provides a personalized view on the system. The display should react when a user approaches the screen and recognize him. The content displayed should adapt based on the person interacting with the screen and the context like which patient the user needs to visit next. There must be a way to transfer information between personal and public devices. It also should offer opportunistic access to relevant information like clinical reports. Favela et al. used the findings to create design scenarios of an extended CHIS that integrates public displays that are able to manage the tasks described above(see figure 6). In figure 4 the different possible functions are made clear, the PublicCalendar displays public events added by the hospital, its staff members and the person watching the screen. The PublicMap application indicates the location of colleagues, available services and the rooms of the patients assigned to the user. The remaining applications depend on the context. In this scenario the user accesses the clinical reports of his next patient and looks at a X-ray image.

The application enables explicit and implicit interaction with the display. Content can be uploaded, downloaded and updated. The user can interact via PDA or handheld table computer. All together the system has a great deal of interaction possibilities.

A 3-min animation that described a potential scenario of the public display in use was created and shown to eleven physicians and two nurses. Afterwards the subjects filled out a questionnaire and were briefly interviewed. Because there is no actual study described the results are highly theoretical. The overall feedback was very positive; the immediate and easy access to medical records was stated to be the most attractive feature. Still there are some privacy concerns when patient data is displayed on a large public screen. It must be assured that only staff members have access to those displays.

In this context the importance of **privacy control** gets visible. But also the implicit interaction eases a lot of the users work and avoids the loss of time through connection process or browsing clinical reports. The application also enhances the collaboration between coworkers by giving the user the possibility to locate the colleague on the screen and contact him, which could lead to an increase of the sense of community.

3 USABILITY FEATURES

Through the analysis of the different applications on public displays several negative and positive attributes could be identified and extracted. In this section the gathered information on usability features will be presented and explained.

3.1 attract attention

The first feature described is the need of an application to attract attention. In modern cities the people's senses are already overloaded with information coming from all the advertising bills, the posters and sometimes even the public displays. If an application wants to stand a chance it has to attract the attention of the passers-by. The analysis of CityWall already showed that, once an application is in use more passers-by notice the display and start to interact with the system. This shows that users or crowds of users can attract the attention of other potential users. But to be noticed in the first place an application needs a feature that stands out. The Polar Defense application for example successfully used an animation to attract the attention of the passers-by. Another way could be the lottery system described in MobiLenin.

3.2 show interaction possibilities

Then the user needs to know how to interact with the system. Therefore instructions must be available and visible. Those could for example consist in an animation shown on the display like it was done in Polar Defense, or simply be the instructions printed next to the screen like in CityWall.

3.3 easy to understand

If a person decides to use an application it must be easy to understand. Often people only use the application for a short period like the time to finish a game round, having explored a few photos or read an article. Therefore the effort put into understanding the application should not be too high, otherwise the person will lose the motivation to use the application. Even complex applications that are used for a longer period of time can be easy to understand if the way of use fits with existing practices or if the functionalities are intuitive.

One example therefore is Dynamo, where the users first had some problems with the carving and dragging principle and needed explicit instructions. The developers overworked the carve and drag principles which led to a way better understanding of the application.

Favela et al. solved this problem in their application for CHISs [11] by using a lot of implicit interaction. So the user does not need to think of how to interact with the system, because the system does most of the interaction work for him.

3.4 avoid social embarrassment

Once an application attracts the attention of the people surrounding, it does not mean the bystanders will necessarily use it. One reason why they could decide not to interact with the display is to avoid social embarrassment. When people feel too exposed, or fear they could fail at a task they prefer not to try to avoid being publicly embarrassed. For example some person did not participate at the game Flashlight Jigsaw because they "didn't want to seem dumb"[4].

A popular device many applications choose as interaction tool, is the mobile phone. This is not only because almost everybody has such a device, but also because it gives the user the possibility to stay anonymous. This is one of the reasons why the mobile phone was chosen as interaction tool for Polar Defense.

3.5 privacy control

Some applications give users the possibility to access private data and information from the public displays [15, 11, 24, 17]. Those applications must discern private data from public data. A person won't use an application when there is the risk that private information is displayed on the public screen so that it can be viewed by everybody surrounding it.

In the CHIS described by Favela [11] concerns about the privacy of patient records were pronounced. Those could be viewed on public

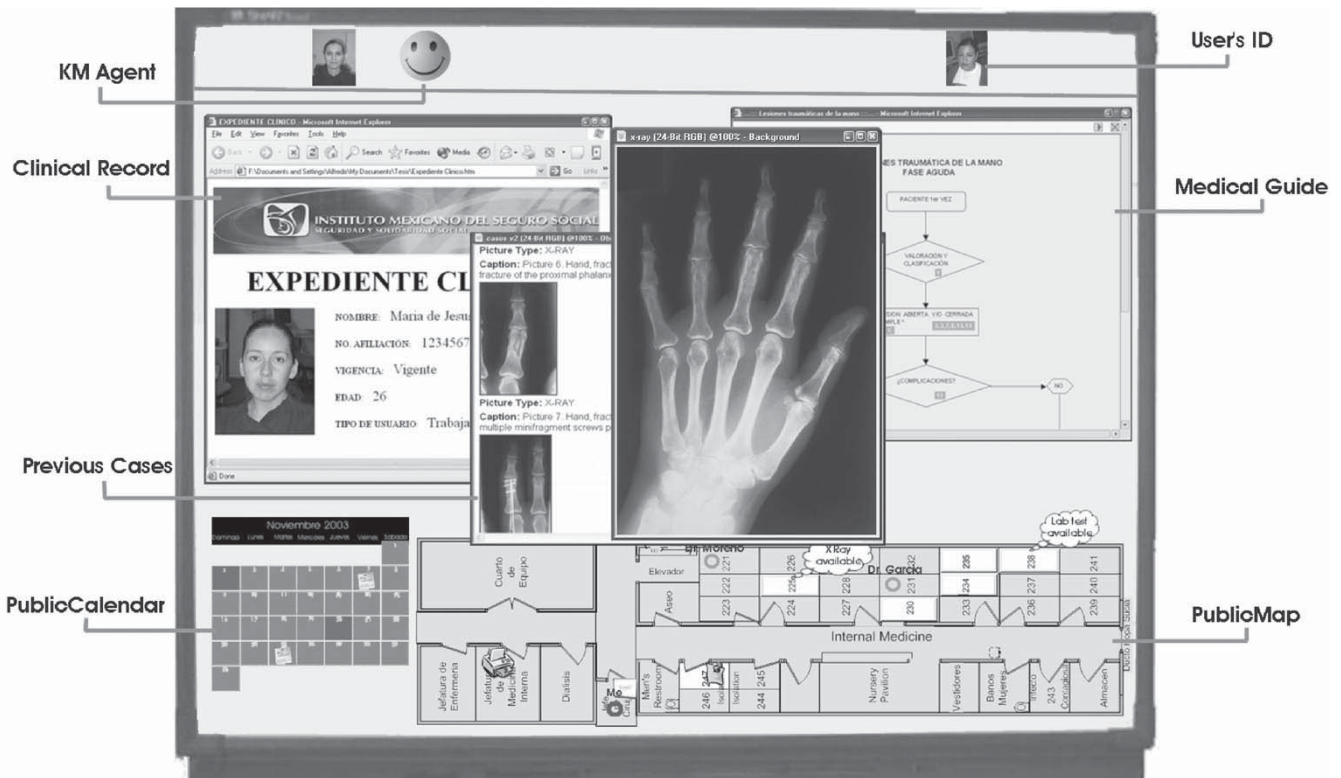


Fig. 6. Personalized view on the public display [11]

displays in the hospital in the same way as public information. Vogel [24] describes a prototype application for interactive public displays where people can access private information. The proximity to the display makes a part of the display switch from a public to a private information display. This part is covered by the body of the user to hide the private information and make it inaccessible to other users. The approach could be one possibility to access private information on public displays.

3.6 conflict management

If the application enables multiple users to simultaneously interact on the same surface conflicts will surely occur. For this reason the application must be designed to avoid conflict situations. If a conflict happens nevertheless the application should offer the possibility to solve it.

The Dynamo system had to deal with many conflicts occurring so that the system was extended and an access control was incorporated. This leads to less conflicts and an easier conflict resolution.

3.7 increase the sense of community

Often the number of persons that use an application declines with the time, people lose interest when the application is not novel anymore. So once a person has used an application it is important to make sure he uses it again. One aspect that can contribute to this is to increase the sense of community. This gives the user the impression that he belongs to a group and motivates him to use the application more regularly.

The results of the observation of the Plasma Poster network during 20 months showed that the community board had managed to build up a group of users that regularly used the application.

4 CONCLUSION

4.1 Summary

The paper shows that there are already many applications developed for interactive public displays. They come from different domains and

offer different interaction possibilities. Applications for collaborative work, to explore images or create conversation opportunities, voting systems, interactive games and community boards have been presented and analyzed.

The results of the analysis made it possible to isolate general features that contribute to a successful application, offering a solution to encountered problems. Regarding those properties an application should be easy to use, show instructions on the different interaction possibilities, control privacy, manage conflicts, attract attention, avoid social embarrassment and enhance the sense of community of a user.

4.2 Discussion and future work

There already has been a lot of work concerning general usability features of applications, which led for example to the usability principles of Schneiderman [21] or the features described by Dix et al. [9]. Those could serve as overall guidelines in the development process of further applications. The characteristics extracted in this paper form more specific guidelines that concentrate on the usability of public applications and that could be used to complement the general principles. They need to be considered in the whole development process, which means in the specification, evaluation and implementation of the application.

Still there are more features that could be identified and further research is needed. A checklist with the different properties could be developed, or some new and more specific features could be isolated depending on the degree of interaction of an application. Similar works even on other types of applications could also be useful. One example is the extraction of usability features for applications on interactive displays for mobile devices. To ease the research more 'best practices' or 'lessons learnt' should be collected to let them inspire the upcoming work.

Even though the applications cover various domains, most of them only have restricted or well known interaction possibilities. Some research on the development of further interaction techniques has already been done like the *Shoot & Copy*-principle [3] where icons on a

display can be selected by photographing them with a mobile phone. The selected information is then directly transferred to the mobile device enabling the user to view, read or listen to the data at anytime. A second possibility is to copy the selected information to a web server and only send the URL to access the information to the user. Other examples are *Sweep* or *Point & Shoot* [1]. Further work could concentrate on finding additional ways of interaction and developing interaction techniques that especially support work on interactive displays in public spaces.

REFERENCES

- [1] R. Ballagas, M. Rohs, and J. G. Sheridan. Sweep and point & shoot: Phonicam-based interactions for large public displays. In *CHI '05*, 2005.
- [2] J. A. Black, J. I. Hong, M. W. Newman, W. K. Edwards, S. Izadi, J. Z. Sedivy, and T. F. Smith. Speakeasy: A platform for interactive public displays, 2002.
- [3] S. Boring, M. Altendorfer, G. Broll, O. Hilliges, and A. Butz. Shoot & copy: Using mobile phones for accessing information on large displays. In *Extended Abstracts of the Ninth International Conference on Ubiquitous Computing (Ubicomp)*, Sept. 2007.
- [4] X. Cao, M. Massimi, and R. Balakrishnan. Flashlight jigsaw: An exploratory study of an ad-hoc multi-player game on public displays. In *CSCW '08: Proceedings of the ACM 2008 conference on Computer supported cooperative work*, pages 77–86, New York, NY, USA, 2008. ACM Press.
- [5] S. Carter, E. Churchill, L. Denoue, J. Helfman, and L. Nelson. Digital graffiti: public annotation of multimedia content. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 1207–1210, New York, NY, USA, 2004. ACM Press.
- [6] E. F. Churchill, L. Nelson, L. Denoue, J. Helfman, and P. Murphy. Sharing multimedia content with interactive public displays: A case study. In *DIS '04: Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques*, pages 7–16, New York, NY, USA, 2004. ACM Press.
- [7] E. F. Churchill, L. Nelson, and G. Hsieh. Caf life in the digital age: Augmenting information flow in a caf-work-entertainment space, 2006.
- [8] N. Day, C. Sas, A. Dix, M. Toma, C. Bevan, and D. Clare. Breaking the campus bubble: Informed, engaged, connected. In *HCI '07: Proceeding of the 21st British HCI Group Annual Conference*, 2007.
- [9] A. Dix, J. Finlay, G. Abowd, and R. Beale.
- [10] E. Eriksson, T. R. Hansen, and A. Lykke-Olesen. Reclaiming public space: designing for public interaction with private devices. In *TEI '07: Proceedings of the 1st international conference on Tangible and embedded interaction*, pages 31–38, New York, NY, USA, 2007. ACM Press.
- [11] J. Favela, M. Rodriguez, A. Preciado, and V. M. Gonzalez. Integrating context-aware public displays into a mobile hospital information system. In *IEEE Transactions on Information Technology in Biomedicine*, pages 279–286, 2004.
- [12] M. Finke, A. Tang, R. Leung, and M. Blackstock. Lessons learned: game design for large public displays. In *DIMEA '08: Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts*, pages 26–33, New York, NY, USA, 2008. ACM Press.
- [13] S. Greenberg and M. Rounding. The notification collage: Posting information to public and personal displays. In *CHI*, pages 515–521. ACM Press, 2001.
- [14] E. M. Huang, D. M. Russell, and A. E. Sue. Im here: public instant messaging on large, shared displays for workgroup interactions. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 279–286, New York, NY, USA, 2004. ACM Press.
- [15] S. Izadi, H. Brignull, T. Rodden, Y. Rogers, and M. Underwood. Dynamo: a public interactive surface supporting the cooperative sharing and exchange of media. In *UIST '03: Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 159–168, New York, NY, USA, 2003. ACM Press.
- [16] J. McCarthy, T. Costa, and E. Liongosari. Unicast, outcast & groupcast: Three steps toward ubiquitous, peripheral displays. pages 332–345. 2001.
- [17] J. F. McCarthy. Using public displays to create conversation opportunities, 2002.
- [18] K. OHara, M. Glancy, and S. Robertshaw. Understanding collective play in an urban screen game. In *CSCW '08: Proceedings of the ACM 2008 conference on Computer supported cooperative work*, pages 67–76, New York, NY, USA, 2008. ACM Press.
- [19] P. Peltonen, E. Kurvinen, A. Salovaara, G. Jacucci, J. E. Tommi Ilmonen, A. Oulasvirta, and P. Saarikko. "it's mine, don't touch!": Interactions at a large multi-touch display in a city centre. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1285–1294, New York, NY, USA, 2008. ACM Press.
- [20] J. Scheible and T. Ojala. Mobilenin combining a multi-track music video, personal mobile phones and a public display into multi-user interactive entertainment. In *MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia*, pages 199–208, New York, NY, USA, 2005. ACM Press.
- [21] B. Schneiderman. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. 2009.
- [22] A. Tang, M. Finke, M. Blackstock, R. Leung, M. Deutscher, G. Tain, C. Giesbrecht, and R. Lea. Designing for bystanders: Reflections on building a public digital forum. In *To Appear in CHI '08: Proceedings of the SIGCHI conference on Human factors in computing systems*, New York, NY, USA, 2008. ACM Press.
- [23] V. H. Tuulos, J. Scheible, and H. Nyholm. Combining web, mobile phones and public displays in large-scale: Manhattan story mashup. In *Fifth International Conference on Pervasive Computing*, 2007.
- [24] D. Vogel and R. Balakrishnan. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 137–146, New York, NY, USA, 2004. ACM Press.
- [25] T. Yamada, J. Shingu, E. Churchill, L. Nelson, J. Helfman, and P. Murphy. Who cares?: reflecting who is reading what on distributed community bulletin boards. In *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*, pages 109–118, New York, NY, USA, 2004. ACM Press.

Private Workspaces in Single Display Groupware and Multiple Display Environments

Dhana Sauernheimer

Abstract— The research domain of collaborative work in interactive surface environments is divided in two subareas. On the one hand, there is a number of different Single Display Groupware systems that basically consist of one jointly used shared surface such as a large wall display or a horizontal tabletop surface. On the other hand, there are Multiple Display Environments which use a combination of some shared interactive surfaces and connected personal devices, for example laptops or PDAs. Investigating how to best support groups of users interacting with a shared display environment includes to explore the handling of private information on these interaction frameworks. For an adequate privacy management it is necessary to perform a separation between public and personal workspaces.

This paper discusses the different techniques to realize this separation. While the abilities of Single Display Groupware are always limited by the singular output channel which makes it impossible to address some private data to one participant exclusively without attaching additional devices to the system, Multiple Display Environments are often characterized by complex synchronization mechanisms, an awareness overload and frequent shifts of attention between devices, which disturbs the natural work flow of the collaboration.

Index Terms—Interactive surfaces, single display groupware, shared workspaces, CSCW, multiple display environments, multi-user interfaces, collaboration

1 INTRODUCTION

Interactive surfaces already offer great opportunities for groups of people collaborating on several occasions. While integrating a variety of different public devices like large touch-sensitive screens or multi-touch tabletop displays with different personal devices (for example laptops, PDAs or USB devices) or personal data of the group participants and embedding them in an ubiquitous computing environment, it is obvious that interactive surfaces have more potential than a simple whiteboard functionality.

The great challenge consists in the development of a framework that allows to perform fundamental collaborative tasks like sharing data across the different platforms, placing it on large public displays and jointly create and modify information [1]. Of course it should be possible to carry out all the provided interactive tasks in a simple and intuitive manner – the usage of the workspace must not cause a disruption of the natural workflow of the group.

Therefore one of the core problems of those frameworks, often referred to as Single Display Groupware (SGD) [16], is the tradeoff between successful collaborative work and the protection of private data. Many environments rely on social protocols [6] or coordination policies [10]. All information is presented on a shared surface itself, but there are implicit personal work areas, for example distinguished by their orientation, position or a label specifying the data owner.

Another solution to differentiate between private and public workspaces is to locate a user's private workspace on his own private device and connect it to the environment. These systems are called Multiple Display Environments (MDEs) [1]. They involve some limitations and transfer problems indeed, but privacy protection will be guaranteed since each participant stays in control of the data published and shared by him at any time in the working process. The collaborators have to decide explicitly about who is allowed to see, manipulate and delete their data. This includes thinking about possible methods and techniques to transfer information between the different work areas.

Discussing both approaches as well as their profit and restrictions,

-
- Dhana Sauernheimer is studying Media Informatics at the University of Munich, Germany, E-mail: sauernheimer@cip.ifi.lmu.de
 - This research paper was written for the Media Informatics Advanced Seminar on Interactive Surfaces, 2009

the structure of this paper is organized as follows. Section 2 gives an overview of different types of Single Display Groupware systems while describing some possible solutions in the field of research. Regarding their limitations, it can be shown that the boundaries between the two approaches are blurring. Section 3 discusses a wide range of Multiple Display Environments in the same way, while the core question of this systems is not how separation of private and public workspaces is accomplished, but how to share information between them and which devices to use for the different tasks and to what extent. Finally, in section 4 the results are summarized and compared in order to give an outlook to some future developments.

2 SINGLE DISPLAY GROUPWARE

Supporting multiple users working simultaneously on a shared surface is a beneficial opportunity to positively influence collaboration, accelerate work proceedings and improve group dynamics. Mandatory requirements are easy-to-learn input and output mechanisms on the one hand, the appropriate handling of personal data on the other. While designing a shared surface environment, the challenge consists in breaking the dominant usage model of "one person per computer" [17] which is not suitable for receiving independent events from more than one mouse or keyboard. Because they cannot be handled separately, all user inputs, regardless of their origin, will be forwarded to one public output channel [17]. This channel is visible to all users: each participant can see and manipulate all information available on the shared interactive surface.

2.1 Usage of coordination policies

In order to avoid the conflict and frustrations resulting from incompatible or interfering input actions executed by multiple users simultaneously, a series of coordination policies [10] was established. These policies are graphical metaphors on the shared surface output that clarify the rights of each participant on the available data by means of social protocols. Thereby, it is able to visually define private workspaces and give the users the sense of being in control of their data as well as more certainty about the way to act in agreement with the social rules of the group.

DiamondSpin [15] is a multiuser environment on a horizontal direct-touch tabletop surface, which is a convenient solution for groups working parallel on visual content in an around-the-table setting. The shared content, for example images, documents or browser windows, is positioned around the table in a circuit.

A set of interaction techniques is provided by the DiamondSpin Tabletop Toolkit [14]. One of the most important coordination techniques is the orientation of the windows. By default, each document is oriented towards the closest outside edge of the tabletop (*see figure 1*). Because it is expected that the participants are sitting around the table and are looking at the content from different perspectives, this scheme guarantees that each user can read the content closest to him, because it faces his edge of the table. Each document can be slid around the table via drag and drop. During this process its orientation is adapted continuously by the toolkit – this is the way the metaphor of “passing the content down to another participant” is realized.



Fig. 1. The user interface of the DiamondSpin tabletop environment. [14]

In order to divide the available space even better in personal work areas of single users, there is a number of further orientation modes and schemes [8]:

- The default mode optimized for a specific number of users is called “Table for Two/Four/Eight”-mode. In this scheme the continuous way of reorientation while moving a document is replaced by a discrete one: each document can only be facing one of the available user edges. This mode results in a clear assignment of a document to a participant – via social protocol it is indisputable that he is the owner of all documents on “his edge” of the table.
- Not only the documents can be dragged around the surface, the background can be directly manipulated, too. This allows the participants to rotate all documents together, as they wanted to switch their workspaces.
- A context-sensitive pop-up menu offers the ability to reorient each document manually. In this scheme, the personal workspaces are deallocated by social protocols: now the surface consists of a collective public workspace, for example suitable for brainstorming sessions.
- A magnetizer on each edge of the tabletop orients all documents to the respective edge. This mode simulates a “one person per computer” situation [17].

Not only the orientation of the a document, but also its size is an important technique to coordinate the privacy level of data in the DiamondSpin environment [10]. Of course users are able to change the size of the visible content via direct manipulation, but there are some automatical schemes, too, for example the “Black Hole Layout” [8], which supports users in working individually on DiamondSpin: the nearer a document at the edge of the tabletop, the bigger it is displayed. Quite contrary to this scheme, in the “Central Focus Layout”

[8] the documents become smaller when positioned near an edge and bigger in the center of the table. This mode is suited for sharing data on DiamondSpin – small documents can be considered as being private.

2.2 Definition of personal areas

Ringel et al. [12] describe and evaluate a series of coordination policies for single display groupware environments using direct manipulation methods on tabletop interfaces. Some of them, especially the “reorient” and “resize” technique, were used in the DiamondSpin framework, as described in section 2.1. In a user study with 30 students Ringel et al. measured the task times and error rates of the participants while sharing and exchanging some documents with four different coordination methods [12]. The results show that the techniques implemented in the DiamondSpin Tabletop Toolkit are not efficient in task time, the error rate was average. The “relocate” technique, which grafically define a specific area of the shared surface as personal workspace for a user, gained the best results by far. It seems to be the easiest metaphor for document sharing to drag an object from a personal workspace into a public one.

UbiTable is an interactive tabletop surface with the purpose to fluidly share, manipulate and exchange data during occasional collaborations at public places [13]. Compared with meetings in front of laptops or most other personal devices, collaborating around a horizontal surface like a tabletop is a more natural way of joint interaction, because “when people sit around a table, a particular table region is in their immediate physical proximity” [14].

Although each participant is able to oversee the whole table and all content on it, the edges closer to a user are more comfortable for him to reach. *UbiTable* uses this effect for defining personal work areas for each participant sitting around the table (*see figure 2*).



Fig. 2. The UbiTable interface with color-coded personal workspaces for two participants and a public area for document sharing. [14]

The “relocate” coordination technique is implemented in the *UbiTable* interface in order to move documents from a personal workspace to the public one – an easy-to-understand metaphor for the process of publishing or sharing data [12] – or to the personal workspace of another participant for exchanging data. Additionally, in the public area in the middle of the tabletop surface each collaborator is able to work actively with the content. Available features are:

- Resize and rotate objects
- Move and rearrange them
- Make copies of file
- Drag a copy of an object to a private laptop which is connected via wireless LAN

- A digital ink function for simple sketches and notes

These features are not only important solutions for many collaborative workflows in meeting sessions, but also implicitly have effects on the privacy level of the data (see section 2.1). While the size, position and orientation of a document only influences the social protocols that let other participants "feel" to be allowed to manipulate the object or not, the possibility to save a copy on a local device really makes the data private.

Documents in a *personal* work area can be viewed by all participants, because the interface can provide only one commonly used output channel for all collaborators, but they can only be accessed by its owner. This is possible due to the capability of the tabletop surface to identify individual users [13]. Therefore, UbiTable is able to distinguish between multiple independent input channels and restrict the access to personal workspaces of other users. A laptop serves as a *private* work area, where documents are only visible to its owner. This separation between privacy and visibility [13] gives the users the ability to control and grant the rights of other participants in a very exact and differentiated manner. Nevertheless, there is no possibility to establish a real *private work area* on the tabletop itself, the connection of a separate device is always required [16].

2.3 User Profiles and Carved Regions

A similar solution to set up personal workspaces on Single Display Groupware is offered by the *Dynamo* system, a framework running on large interactive surfaces which may be composed of one or more displays like screens or tabletops. Like UbiTable, *Dynamo* supports groups of people in spontaneous face-to-face collaborations to easily share, display and exchange different types of digital media sources [6]. Users can either connect remotely to the multi-cursor environment via a personal computer, PDA or laptop or attach USB input devices like cameras, MP3 players and removable disks through so called interaction points [2]. Each interaction point is represented on the surface by a color-coded cursor which can be controlled by a connected device with the ability to provide mouse and text input [6]. The device is treated as a second display, so users can simply drag all supported content types from their device to the surface to share them on the public *Dynamo* workspace. They can also leave data packages (called *media parcels* [2]) on the surface to exchange them with other participants.

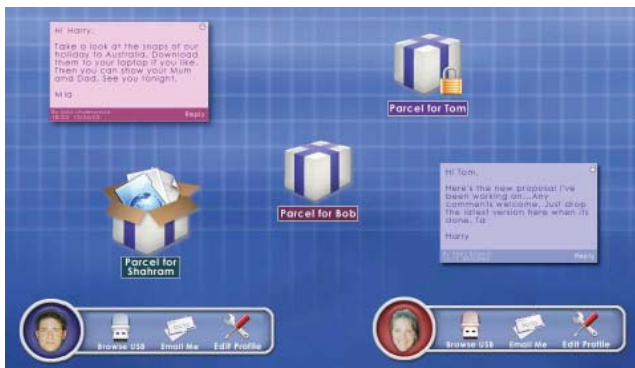


Fig. 3. The *Dynamo* user interface with two connected participants, identified by their personal palettes (at the bottom), some color-coded notes and media parcels. [6]

One of the core ideas of the *Dynamo* system is the persistent identification of users via profile data (containing their name, contact information and a picture). Although it is possible to use the surface with a guest profile identification brings many advantages to the users. Once connected, a *personal palette* with their profile picture icon is shown at the bottom of the surface (see figure 3). These palettes allow the user to leave personal notes identified by his profile color, and provides each collaborator the ability to leave private media parcels that

are addressed to him. A private media parcel is displayed on the surface like a public one, but only the identified receiver is able to open it or drag it on the private workspace on his connected device [6].

Dynamo also provides an interesting concept to realize personal workspaces on the shared display. Identified users can "carve off" a part of the surface and set the access rights for this region [6]. Due to the limited output options of Single Display Groupware, each *carved region* is always visible to all users, indeed. But only the owner can manipulate the content that is placed inside. By means of dragging the icon on another user's personal palette in the created carved region (see figure 4), the owner can invite participants to join his area. So they receive the full access rights, too, and a spontaneous collaborative working session can be established among them [2].



Fig. 4. A carved region with two participants (right), labelled by the key on the top. Only the owner and invited users can work in this area. [6]

Even asynchronous collaborations are supported by the *Dynamo* system. Carved regions as well as public or private media parcels can be stored persistently on the surface [6]. So a meeting can be suspended and continued at a later time. Furthermore, it is possible to leave data for users who are currently not connected to the system. It will be displayed on the surface as a media parcel for the specific user and be ready for collection at any time.

2.4 Limitations and problems

All the Single Display Groupware systems introduced in this section try to explore and implement reasonable approaches to make it possible for multiple users to work simultaneously on a shared interactive surface environment. Table 1 shows a short classification of the three frameworks in terms of their I/O and workspace management.

Table 1. Single Display Groupware

	DiamondSpin	UbiTable	Dynamo
Input via	direct touch	direct touch	mouse and keyboard
Output via	tabletop	tabletop	any combination of large surfaces
Public space	the whole surface	in the middle of the surface	public regions
Personal space	by social protocols	side areas on the surface	carved regions
Private space	n/a	external device	external device

There is one critical limitation that all Single Display Groupware systems have in common. Because there is only one common output channels for all participants together, there is no possibility to establish private areas on the shared display, which are only visible to their owner or participants with the permission to view their content. UbiTable and *Dynamo* solve this problem by differentiating between

private and *personal* areas. Personal workspaces are located on the shared surface – their content is visible to everyone, but only manipulable to users with permission. The private workspace, that is only visible to its owner, has to be swapped out to an external device, for example the user’s laptop.

DiamondSpin does not provide any private work areas at all. Even a personal workspace is not defined explicitly but with the aid of a set of coordination policies that rely on social protocols, which are just rough guidelines of action. The problem with these metaphors is that wrong actions (for example a user tries to access private data of a collaborator unauthorized) cannot be prohibited automatically but via reactions of other users. The right handling of the system and the correspondent coordination policies have to be learned by trial and error, which may cause uncertainty among both the data owner, who feels insecure concerning the insufficient control over his information [6], and his collaborators, who can never be sure about the privacy level of the provided content.

Another important point is the real-estate problem [17] which often occurs when multiple users share a limited space, as it is the case in Single Display Groupware. Every user needs individual widgets on the shared surface that offer permanent feedback about all of his actions. The disorder resulting when too many users participate can easily cause disorientation and a fatal awareness overload [16], particularly if there are many duplicate elements such as telepointers and personal palettes in a multi-cursor environment like Dynamo. Although in this framework the overload problem is solved by the capability to extend the surface by attaching further displays at any time, direct touch surfaces like tabletops or interactive wall displays are more qualified in this regard. Moreover, they have other critical disadvantages: firstly the maximum number of participants is inherently restricted by the physical space around the surface, secondly environments like UbiTable always can only run properly on very special and expensive tabletop surfaces that support multiple user inputs and are able to keep them apart.

3 MULTIPLE DISPLAY ENVIRONMENTS

In order to find a way out of the real-estate problems and privacy limitations often occurring in Single Display Groupware systems, an obvious solution is to create environments that use more than one interactive surface for fulfilling the collaborative tasks.

Similar to the possibility to attach external devices to UbiTable and Dynamo to establish private workspaces, the Multiple Display Environments (MDEs) [1] that will be described in the following section are composed of a number of shared interactive surfaces like large wall displays or tabletops on the one hand, personal devices like laptops on the other. The difference to Single Display Groupware is the intensified working with the connected personal devices. They are not only used as a storage for private data that needs to be invisible to other participants, but also as real workstations during the collaboration process. The key challenge of the infrastructure consists in the need to support cross-surface interaction. Some interesting solutions are described below.

3.1 Hypermedia Models

An early approach is offered by the *DOLPHIN* environment. An interactive electronic whiteboard, a Xerox LiveBoard, is used in *DOLPHIN* as the hardware basis and a central public workspace in the meeting room [9]. Personal computers can be connected via Ethernet to the system – either in the local meeting room or as external partners remotely. So *DOLPHIN* is able to support both computer-supported face-to-face meetings and remote sessions.

The connected workstations operate as private workspaces for the participants. The *DOLPHIN* clients consist of an editor widget following the WYSIWIS (What You See Is What I See) principle [18] that allows the users to view a duplicate version of the current presentation on the whiteboard and edit its contents on their personal computers with mouse and keyboard. Therefore, a second widget is located on the client, the private work area, which is not displayed on the LiveBoard and hence not visible to other participants. It can be used to

locally generate new content and simply drag it to the public space after finishing. Analogously, data from the public area can be copied to the private one for editing and updating them. Of course the whole other functionality of the personal computers can be used private by the participants during the session, for example a mail client, a web browser or a calendar.

A core question while developing the *DOLPHIN* framework was how to deal with the tradeoff between the freedom of users to be able to present as much different document types as possible on the whiteboard and the technical capabilities to support them on all connected devices [18]. Finally, a *hypermedia data model* was adopted. This model implies that all data which is to be displayed on the whiteboard is represented by a hypermedia node. These nodes can contain different types of links to other nodes, so relationships between them and parts of them can be created [9].

The following types of hypermedia objects are supported by the *DOLPHIN* Cooperative Hypermedia Engine Client [18]:

- Scribbles – pen based handwritings produced on the LiveBoard, drawings or different figure (circles, boxes, tables, arrows...)
- Text – editable ASCII objects
- Images – raster graphics image objects

It is obvious that the heavily limited abilities of *DOLPHIN* to include different document types in the hypermedia data model – for example rich text documents, web pages, formatted code, or even music and video file – makes the system unqualified for productive collaborative meetings nowadays. Even similar environments like the Notification Collage [5], which are able to publish more advanced documents such as video notes, desktop snapshots and web page thumbnails, share those restrictions.

Because participants have to manage the whiteboard’s hypermedia structure themselves, they will be easily distracted from their actual working tasks. This can result in a disruption of the natural workflow of the meeting. In addition, users are often focused on their own personal workstation and not on the central public workspace, the LiveBoard, because they will often create and edit nodes locally on their private workspaces and present their finished ideas instead of letting them grow together.

The lessons learned from this early MDE solution are applied in further developments realizing multi-surface environments with trying not to restrict user’s freedom and to enable a more natural way of computer-supported collaborative work. This can be achieved with a completely different approach: Instead of transferring some predefined types of hypermedia objects and store them persistently on a cooperative hypermedia engine server [19], the idea is to transmit the whole screen image of the personal device to the public workspace. Thus, a joint focus for all participants is established without restricting their individual working process.

3.2 Screen Sharing

The *WeSpace* is a Multiple Display Environment that uses screen sharing via virtual network computing (VNC) [11] to transmit personal workspaces one-to-one to a shared public surface [20]. The environment consists of a large high-resolution wall display and a direct-touch tabletop. Users have to bring laptops to the collaborative work session, sit down around the table and connect to the system with a client installed on their devices. The screen images will be rendered on both public surfaces with 15 frames per seconds using OpenGL [19].

Of course the *WeSpace* supports the display of multiple user screens at the same time. With the tabletop interface the participants can arrange and resize the screens manually via drag and drop, additionally there is a control panel that allows to set automatical display priorities for all screens that regulate their privacy status [19]: the screen images can be set to public or private (invisible) mode. They can also be set ‘important’, then they will appear extra large on the wall display, for example for a user’s presentation. In all cases, each collaborator stays in full control over the privacy and visibility level of his data.

A contributing factor may also be the wide choice of interaction methods: each screen display is controllable directly by mouse and keyboard on the laptops or indirectly by direct touch on the tabletop surface. For simulating mouse commands on the tabletop, users have to activate the separate view [20] that enters a full screen mode of the current laptop screen image. Now direct touch input will be forwarded to the laptop. The view is called separate, because the wall surface is not displayed synchronously to the table until the user gets back to the synchronized view. An application named LivOlay running on the WeSpace environment makes the arrangement of screen images even more comfortable: it enables the so called overlapped view [19]. The z-index of the widgets as well as their transparency can be simply set with some simple gestures on the tabletop.



Fig. 5. Astrophysicists meeting in the WeSpace. [19]

The great advantage of the WeSpace compared to interactive whiteboard environments as described in section 3.1 consists in the support of natural individual work which can be easily integrated in collaborative meeting sessions. Users do not have to care about the system's functionality because their own workspace, which is familiar to them, is mapped on the shared surface. They can run their own applications and display each type of information and document.

By using the tabletop as a joint focus workstation (see figure 5), there will be no separation and encapsulation of single participants who only pay attention to their private devices and the collaboration will not be disturbed and can proceed more fluent. Only keyboard input is not possible with the tabletop interface and must be done directly on the laptops.

However, the WeSpace is not applicable for easy document exchanging between the collaborators as supported and focused in the presented Single Display Groupware systems. Additionally, there can occur an overload problem when too many users are connected: not all their displays can be displayed in an appropriate size on the shared surface and because of limitations of the separate view parallel working on more than one screen via the tabletop interface is not possible.

Another MDE approach, that tries to concentrate more on these tasks, is described in the following section. In this case, the idea is not to share hypermedia objects or whole screen images, but off-the-shelf applications between personal and shared devices without modification.

3.3 Application Sharing

The *IMPROMPTU* interaction framework [1] was developed in order to improve collaborative problem-solving, for example in groupbased face-to-face software engineering. In a number of field studies [1], Biehl et al. tried to find out how Multiple Display Environments are used in authentic work sessions for real activities.

The resulting idea was to share task information across devices via off-the-shelf applications. This concept was already implemented in the *iRoom* environment [7] by sharing URLs between devices, that describe the information content. Unfortunately, the installation of all the

used applications on every connected device is required in this solution [7], so *IMPROMPTU* is aimed at allowing any off-the-shelf application to be shared without modification on interactive public surfaces and enabling parallel multiple user input on it. This includes maintaining the whole interaction context of the application [1], for example an ancillary widget providing additional information, and, of course, options for managing privacy issues.

The *IMPROMPTU* user interface on the personal devices consists of three components. On the left side of the screen, the *Collaborator Bar* is displayed [1]. It shows a list of all connected participants with their names and an image icon. Each collaborator profile can be expanded per mouse click. In this mode, an overview of all application windows shared by this user shows up as thumbnails (see figure 6). These thumbnails can be dragged and dropped on the desktop to create a replication.

For each connected shared surface, there is one *Shared Screen Dock* on the top of the user interface [1]. It also shows window thumbnails for all applications replicated and shared on the surface and offers the ability to drag widgets on the surface or back. While dragging a document to the Shared Screen Dock, it will expand, so the position and size of the window on the surface can be easily set manually by the user.



Fig. 6. An expanded user view on the Collaborator Bar. All applications provided by the participant Jake in share or show mode are displayed as thumbnail with an icon indicating the document type. A replication of each window can be created by dragging the thumbnail to the desktop. [1]

At last, there is a component to let the users configure privacy settings for their applications. It is called *Collaboration Control* [1] and allows the application owners to differentiate between three privacy modes, as used in many Single Display Groupware systems (see section 2). Applications in *share* mode are public, that means that all participants can view them and edit their content. Applications in *show* mode are personal. They can be viewed only, but no interaction is possible for collaborators. Private data, that is invisible to other users, is set in *Do not Show or Share* mode and will not even appear in the Collaborator Bar of the owner.

3.4 Limitations and problems

Multiple Display Environments have more abilities than Single Display Groupware to improve and manage privacy while using shared surfaces collaboratively, because a part of the work takes place on private devices. A separation between public and private work areas is given by the hardware settings anyway. However, the parallel usage of different types of devices can be a challenge since adequate transfer techniques must be developed and the focus shift of the users between public and private displays needs to be considered in the design process. As a result, the presented Multiple Display Environments

strongly focus on different user tasks and activities (while disregarding other ones) and therefore follow completely different approaches. A classification is illustrated in table 2.

Table 2. Multiple Display Environments

	DOLPHIN	WeSpace	IMPROMPTU
Input via	mouse, keyboard, pen	direct touch, mouse, keyboard	mouse, keyboard
Public output	LiveBoard, public space on PC	wall display, tabletop	customizable (any public surface can be connected)
Private space	PC	PC or laptop	PC or laptop
Privacy management	remove single hypermedia objects	set screen as 'private'	set single applications to 'do not show or share'
Main attention focus	PC	tabletop and wall display	PC or laptop
Document exchanging	only predefine hypermedia types	n/a	via application sharing

DOLPHIN's possibilities are limited to an electronic whiteboard functionality for multiple users. It may be used to support simple brainstorming sessions, draw sketches or to jointly structure information. Because hypermedia nodes and links are transferred to and stored on the server, fluid parallel interaction with the system is possible and the exchange of the shared information between participants can be easily performed. Unfortunately, the sharing of hypermedia objects always involves the restriction to a predefined set of document types, which makes the system unsuitable for broader deployments.

Another great disadvantage, that can occur on both systems DOLPHIN and IMPROMPTU, is the disturbed workflow because of the encapsulation of participants who are focused on their personal computers most of the time, which can be hindering in collaborative working processes. There's the risk that users increasingly work alone on their private workspaces and only post their results to the public workspace on the LiveBoard after finishing

Quite contrary to DOLPHIN, the WeSpace environment aims at the ability to publish any type of documents and applications on interactive surfaces and align the main attention focus of all participants to a commonly used input device, the tabletop surface. The system seems to be a good solution for presentations and collaborative work with a single private device as basis, indeed, but the user interface on the tabletop provides no support for simultaneous work on different laptops. In this case, the collaboration process will be interrupted and each user will get back to individual work on his private device. Moreover, there are no integrated mechanisms to exchange information between the different devices.

IMPROMPTU covers all those options and guarantees a great flexibility because several devices, both private workstations and public interactive displays, can be connected to the system, addressed and controlled by the participants and used to share and jointly work on any off-the-shelf application while its privacy settings are fully manageable by its owners. Some complex mechanism for synchronization mechanism between the replicated windows on the different devices [1]. The real-estate problem can be easily avoided by connecting an additional public surface. However, too much extension can always cause an awareness overload, especially if the users are required to perform many attention shifts between the different surfaces.

4 CONCLUSIONS AND FURTHER DEVELOPMENTS

Both approaches to support collaborative work by using interactive surface environments sometimes reach their limits when a separation of private and public workspaces is required.

Single Display Groupware systems on the one hand seem to allow the participants more natural and fluid communications because

the main attention focus lies on one shared interactive surface. But these solutions often cannot completely get over the model "one person per computer" in an effective way [17] which causes an insufficient separation of private and public work areas. Private data can not be turned invisible without extending the system by connecting some private devices. Multiple Display Environments on the other hand can provide a better privacy management since private and public spaces are physically separated [4], but they restrict the collaboration process because of encapsulation of users and attention shift problems [16]. A challenge consists in finding an adequate combination of the two approaches.

A concept often discussed in the field of Single Display Groupware is the *spatial multiplexing* [3] of interactive shared surfaces. This technique is often used to improve the mapping of user actions, for example several types of direct-touch input gestures, to commands. The surface's user interface can be graphically divided in regions for different tasks, so called *modal spaces*. When a document is moved to this region, the associated command mode is activated automatically [3]. Moreover, by differentiating between different input types (for example one-finger touch - two-finger touch [3]), the set of available action modes that deserve no further declarations or menu selections can be easily expanded, assuming that the gesture design is simple to memorize.

For future work, it may be beneficial to investigate how a separation of public and personal workspaces in Single Display Groupware can be improved by mapping privacy levels to the different input actions supported by the interactive surface [3]. Therefore, the feedback output that is addressed to a single user but displayed publicly on the jointly used output channel of the surface can be reduced. As a result, the real-estate problem (see section 2.4) will decrease and transitions between the public, personal and private modes will be simplified

Another valuable idea is the development of so called *Single Display Privacyware* [16] systems. They are an extended approach of Single Display Groupware and are based on a simple principle: without losing the interaction context given by the public shared output channel of the interactive surface, the systems provide users the opportunity to receive additional private output, that is individualized for their personal benefit and invisible to other participants [17] (see figure 7).

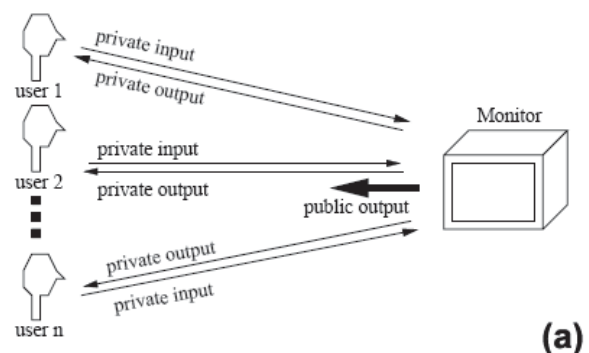


Fig. 7. The basic principle of Single Display Privacyware. While classic SDG only provides one jointly used public output channel for all users, these systems also include private individual output, dependent of input actions of a single user. [16]

This can be achieved for example by the usage of CrystalEye glasses. Instead of providing a stereo view for a single user, an altered shuttering sequence [17] enables individualized views for two users: The first pair of glasses is modified so that both lenses are opened during odd-numbered refresh display frames and closed during the even-numbered frames – the second pair is altered vice versa [16]. So the participants see different refresh display frames on the shared surface. They can be used as separated output channels and

provide a real private workspace on a public screen. An example application scenario is the supply of hints during a multiplayer games (see figure 8).



Fig. 8. Single Display Privacyware (example): During a multi-user football game, each player receives individual moving hints. [16]

Of course, these approaches are strongly limited, yet – Crystal-Eye glasses are not comfortable enough in most cases and the principle does not work in larger groups of users. Also the usage of spatial multiplexing, modal spaces and gesture mapping in order to support collaborative work is barely explored. It would be valuable to investigate further solutions in future work and enhance combinations of Single Display Groupware and Multiple Display Environments to tap the full potential of privacy management in collaborative work.

REFERENCES

- [1] J. T. Biehl, W. T. Baker, B. P. Bailey, D. S. Tan, K. M. Inkpen, and M. Czerwinski. Impromptu: A new interaction framework for supporting collaboration in multiple display environments and its field evaluation for co-located software development. In *Proceedings of ACM CHI 2008 Conference on Human Factors in Computing Systems*, pages 353–362, New York, NY, USA, 2008. ACM Press.
- [2] H. Brignull, S. Izadi, G. Fitzpatrick, Y. Rogers, and T. Rodden. The introduction of a shared interactive surface into a communal space. In *CSCW '04: Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 49–58, New York, NY, USA, 2004. ACM Press.
- [3] K. Everitt, C. Shen, K. Ryall, and C. Forlines. Modal spaces: spatial multiplexing to mediate direct-touch input on large displays. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, pages 1359–1362, New York, NY, USA, 2005. ACM.
- [4] S. Greenberg, M. Boyle, and J. Laberge. Pdas and shared public displays: Making personal information public, and public information personal. Technical report, Personal Technologies, 1999.
- [5] S. Greenberg and M. Rounding. The notification collage: Posting information to public and personal displays. In *CHI Letters*, pages 515–521, 2001.
- [6] S. Izadi, H. Brignull, T. Rodden, Y. Rogers, and M. Underwood. Dynamo: A public interactive surface supporting the cooperative sharing and exchange of media. In *UIST '03: Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 159–168, New York, NY, USA, 2003. ACM Press.
- [7] B. Johanson, A. Fox, and T. Winograd. The interactive workspaces project: Experiences with ubiquitous computing rooms. *IEEE Pervasive Computing*, pages 67–74, 2002.
- [8] M. E. R. Laboratories. Merl - diamondspin. <http://www.merl.com/projects/diamondspin/>, 2008. visited 11.06.2009.
- [9] G. Mark, J. M. Haake, and N. A. Streitz. Hypermedia structures and the division of labor in meeting room collaboration. In *CSCW '96: Proceedings of the 1996 ACM conference on Computer supported cooperative work*, pages 170–179, New York, NY, USA, 1996. ACM.
- [10] M. R. Morris, K. Ryall, C. Shen, C. Forlines, and F. Vernier. Beyond social protocols: Multi-user coordination policies for co-located groupware. In *Computer Supported Cooperative Work, Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 262–265, New York, NY, USA, 2004. ACM Press.
- [11] T. Richardson, Q. Stafford-Fraser, K. R. Wood, and A. Hopper. Virtual network computing. *IEEE Internet Computing*, 2(1):33–38, 1998.
- [12] M. Ringel, K. Ryall, C. Shen, C. Forlines, and F. Vernier. Release, relocate, reorient, resize: fluid techniques for document sharing on multi-user interactive tables. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 1441–1444, New York, NY, USA, 2004. ACM.
- [13] C. Shen, K. Everitt, and K. Ryall. Ubitable: Impromptu face-to-face collaboration on horizontal interactive surfaces. In *In Proc. UbiComp 2003*, pages 281–288, 2003.
- [14] C. Shen, K. Ryall, C. Forlines, A. Esenther, F. D. Vernier, K. Everitt, M. Wu, D. Wigdor, M. R. Morris, M. Hancock, and E. Tse. Informing the design of direct-touch tabletops. *IEEE Comput. Graph. Appl.*, 26(5):36–46, 2006.
- [15] C. Shen, F. D. Vernier, C. Forlines, and M. Ringel. Diamondspin: an extensible toolkit for around-the-table interaction. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 167–174, New York, NY, USA, 2004. ACM.
- [16] G. B. D. Shoemaker. Single display privacyware: Augmenting public displays with private information, November 2000.
- [17] G. B. D. Shoemaker and K. M. Inkpen. Single display privacyware: Augmenting public displays with private information. In *Conference on Human Factors in Computing Systems - Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 522–529, New York, NY, USA, 2001. ACM Press.
- [18] N. A. Streitz, J. Geissler, J. M. Haake, and J. Hol. Dolphin: integrated meeting support across local and remote desktop environments and liveboards. In *CSCW '94: Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pages 345–358, New York, NY, USA, 1994. ACM.
- [19] D. Wigdor, H. Jiang, C. Forlines, M. Borkin, and C. Shen. Wespace: the design development and deployment of a walk-up and share multi-surface visual collaboration system. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 1237–1246, New York, NY, USA, 2009. ACM.
- [20] D. Wigdor, H. Jiang, C. Forlines, M. Borkin, and C. Shen. Wespace: the design development and deployment of a walk-up and share multi-surface visual collaboration system. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 1237–1246, New York, NY, USA, 2009. ACM.

Hybrid Interaction on Interactive Surfaces

Stefan Grabs

Abstract— Interactive Surfaces offer a new type of direct manipulation of digital information. These multi-touch interfaces support inputs by simply using fingers. But some aspects of control were left behind. These systems do not support a tactile feedback or a precision which mouse and keyboard have. This paper gives a review of a human-computer interaction system that uses both this natural - multi-touch - and physical user inputs to directly manipulate digital information on interactive surfaces. These hybrid interactions are an extension of pure touch interfaces and a combination with parts of tangible user interfaces to resolve the lacks of former manipulation by just using the finger. Users generate a connection between the material and digital world by simply putting the artifacts on the surface and get back a tactile feedback for example with using special physical objects to get a better understanding of the mean of their actions. Finally the paper will state the benefits of using hybrid interaction systems regarding different user groups. And in addition it shows that a classification of the used objects can easily be drawn with the result of developing a drawer system of different physical artifacts.

Index Terms—hybrid interaction, interactive surface, tangible interface, natural interface, physical manipulation, social feedback, physical widget

1 INTRODUCTION

Connecting mobile or other devices to a computer is usually a lot of work. Usually the user has to use a cable or set up a whole new network to interact with the specific artifact. These processes take time and constrain the usability which had become an important factor over the past years. Another aspect is the use of mouse and keyboard which became the most popular ways of controlling computers but the user needs a special learning process and training for utilizing these devices because they include no natural gestures [35]. Previous work dealt



Fig. 1. Hybrid interaction on an interactive surface [36].

with these problems, and researchers created new human-computer interfaces to improve the usability. This so called touch interfaces; also known as touch or multi-touch systems, support direct manipulations by the user by simply using the finger or fingers to make inputs. But with the development of these systems further points of interests were revealed: The user is able to use natural gesture for controlling but the

problem of using external devices and at first the use of mobile devices remained. So changing digital information via physical objects and creating a link between real and virtual realm became another field of research. One of the first and a widely recognized concepts in that field was the "Marble Answering Machine" [9] described by Durrell Bishop in 1992. It is an answering machine which uses marbles for representation of messages which had been saved. The user can pick them up and put them into the player queue for playing them back. Furthermore the marbles can be put into another queue to directly dial the number of the caller. These kinds of systems are controlled just by using physical artifacts, so their name was derived by their interface ability: tangible or graspable user interface. These two types of interfaces "touch and tangible" create a new class of interactions, the hybrid one (figure 1). These interactions contain the usability of natural gestures and the comforts of using objects of the material realm which are linked to digital data by putting them onto the interactive surface [7]. For example, the reacTable is a project that supports both touches and inputs via physical objects placed onto the screen to provide live music. With the basics of interactive surfaces and different user interfaces it is important to classify hybrid objects to get a better understanding of the functionality and usability of hybrid interfaces. By giving examples and showing the ease-of-use and functionality of hybrid interaction on interactive surfaces, the paper demonstrates the extensive change of computer interactions by simple diversify the interface. Further another aspect is contained: in what ways hybrid objects could be classified. That leads to important questions like location, way of manipulation, tracking etc. into consideration. Furthermore the difference between the digital and physical realm itself plays an important role. For example, the benefits of using physical artifacts instead of 2D animations which are only visible on the screen are an often discussed topic. [29] In this paper I give an overview of hybrid interaction on interactive surfaces as well as I state the difference of using graspable objects together with touches instead of using other interfaces and the importance of these facts for the user.

2 DEFINITION & TERMS

It is very important to know terms and definitions while discussing hybrid interaction on interactive surfaces. This paper will state a few of them to give a better introduction about the whole topic and to show the diversified understanding of these terms and definitions regarding diverse special interest groups or researchers.

2.1 Interactive Surfaces (IS)

Interaction is the ability to directly manipulate the upcoming behavior of a system and the interplay between actor and reactor. Derived from this we can describe an interactive surfaces as physical surface

- Stefan Grabs is studying Media Informatics at the University of Munich, Germany, E-mail: stefan.grabs@campus.lmu.de, Matrikelnummer: 6025985
- This research paper was written for the Media Informatics Advanced Seminar on Interactive Surfaces, 2009

with the ability of a directly manipulation. These systems can even be found in nature as we take a closer look: An example is skin (regarding this, the user would be the environment). So every surface that gives a feedback in any ways, whether tactile or not, and has a direct manipulation control to change the system's state can be described as an interactive surface. An interactive surface in the field of computer technology stand for a form of human-computer interacting [29, 28]. These systems contain lots of interaction possibilities. Almost always digital data are displayed by projectors or screens and the user can easily interact with them by using physical artifacts or body parts, like his or her fingers. Furthermore there is no actually need of a screen. The screen does serve just as a visual feedback for the user. So the interactive surface stands for a direct manipulation of a virtual environment. In addition it forms the fundament for ongoing research on other ways of input and system control. As mentioned above new interfaces developed from the thought of interactive control: The touchable and the tangible user interface are the two most important ones for this paper because these two lead to the hybrid interactions discussed here.

2.2 Touch Interfaces (TI)

This type of interfaces, the natural one, was one of the first possibilities to control computer systems by using fingers or other natural gestures [8]. Here the word natural has two different meanings first it stands for the way of movements: Natural congenital gestures with no need of learning them. A possible example for this would be the pointing on things appearing on the screen in contrast to system control by mouse. The second meaning represents the way of input: Not just the gestures are natural, the inputs too. Regarding the same example; the click on the object that we found on the screen is not realized with a pen or another physical object, it is simply realized by a touching it. The first touch interfaces embedded a pure touch-screen which supported just one simple input at the same time but in ongoing researches, these simple systems were developed further to multi-touch and then to multi-user systems which supports more natural inputs of different users at the same time [10, 8]. So the control was extended and the possibilities for users expended to a collaborative playground, for example Microsoft Surface supports up to 52 finger inputs at the same time [1]. Comparing the old graphical user interface which is realized with mouse and keyboard control with the touch interface that supports a multi-touch screens that is much more intuitive, contextual and evocative [3]. The only problem or disadvantage with the new interface is that the user has no longer tactile feedback. The lost physical feedback while pushing buttons or scrolling raises the lack of fluidly interaction and precise control [34].

2.3 Tangible user interfaces (TUI)

A tangible user interface also known as TUI is an interface that supports physical artifacts as representation and control elements [21]. The user is able to make inputs by simply putting those objects onto the surface. The system recognizes them by using one of a various number of tracking methods and creates a digital response, for example enable a digital shadow. This digital or virtual shadow is the digital surrounding which appears when objects are connected with the underlying surface. It links object and digital information. Furthermore there are advantages of graspable interfaces in social network, too: Collaboration and the sharing of data become easier and the whole work lead to a much more personal way. Beyond this the use of these artifacts is a passive one after enabling the virtual reaction the objects need no further wires, batteries or other equipment as Brave and Ishii state it [6]. It is possible to say, like Mazalek et al [18] did, that tangible interfaces are rooted in our physical surroundings, employing objects, surfaces and spaces as embodiments of digital information. Some projects in this field of research are: Algoblock [27], a brick system of physical objects to create a program, ToolStone [23], a tool for the non-dominating hand as support for copy and paste or other actions and Urp [32] which analyses tracked objects and generates certain environmental conditions, like a shadow on the screen that belongs to the shadow of a certain day time, to support urban planning.

2.4 Hybrid interaction

Hybrid interactions on interactive surfaces are a mixture between the above explained interfaces. Hybrid stands for a combination of two or more things, in our case of two things: The input via fingers or other body parts and the input through physical objects while the artifacts create a link between themselves and the digital data they referring to [28]. These objects are often called hybrid widgets or just widgets, because they represent a widget of the user interface [7]. Hybrid interactions are always associated with an interactive surface containing a screen. With this new form of interactivity researchers tried to fix the problems which natural and tangible user interfaces raised. Now the user is able to control the system using natural gestures and with the help of artifacts which retrieve a tactile feedback. An example are SLAP Widgets [34] which are transparent widgets composed of silicone and acrylic. These widgets are put onto the screen to give users physical feedback. Beyond that the usage of hybrid widgets opens new ways of interactivity and collaboration. The artifacts serve as support for making modifications on digital data much more easy like the example SLAP or PhotoHelix [7] demonstrate. Researchers see this concept and the resulting widgets as a compromise between graphical and tangible user interfaces. Butz et al describe hybrid widgets in [7] as:

recognizable and reusable interface objects, but also provide a physical handle. Interaction with them takes place in a very physical way, by manipulating the handle with the hand, but they also contain a digital part, which is closely coupled to the physical handle. [...] This combination of a physical and a digital artifact provides a haptic quality to interaction with the screen content, but at the same time uses the digital shadow to provide functionality which goes far beyond the capabilities of the physical object alone. So these hybrid ways of interaction establishes a further and at least a richer diversity of interactions on interactive surfaces.

3 COMPONENTS OF AN INTERACTIVE TABLETOP

An interactive tabletop is built up with a table. That table can look very similar to regular tables as for example the Bluetable [37] or have a special form. Almost always form is generated by usefulness and because interactive surfaces often serve as collaboration system the table form was chosen to support multi-user inputs in a proper way [26]. The important thing is the top surface of the construction. It represents the human-computer interface. Referring to the type of that interface it supports different kinds of input objects. Furthermore the tracking of these inputs is an important part for setting up an interactive system. Tracking means the recognition of where and in addition what kind of object the inputs made. This paper presents three types of tracking [30]:

- Optical Tracking Technique which uses a camera which supervises the surfaces and the computer system of the construction evaluates these pictures to send back a digital response.
- Acoustic Tracking Technique that utilizes microphones in the corners of the surfaces which recording the input noises and the system processes them for tracking.
- Capacitive Tracking Technique which uses the electric potential coupling of transmitter and antenna, the system identifies whether voltage lies on or not and tracks the input point by using an intern grid.

3.1 Optical Tracking Technique

This tracking method uses one or more cameras to take pictures of the workspace. The system processes these pictures with the help of a graphics editing program and calculates the points of input into coordinates. Different techniques have been established yet. One of the well-known methods is frustrated total internal reflection also known as FTIR [15, 12] which can be differentiated into two types. Both of

them use infrared light which is invisible for the human eye. A projector displays the information. First there is the strategy to put parts of the systems, infrared camera and projector, above the surface. The infrared light covers the whole screen and the user can operate simply by touching on the surface. When a user interacts with the product the infrared light beams are refracted and the system recognizes where an action took place. The gathered information is evaluated and new information is displayed by the projector. This technique is often used for interactive tabletops, an example is the Tangible Viewpoints system [18] or the BlueTable [37]. The second method is to keep the camera and the projector below the screen. (as figure 2 shows it). The

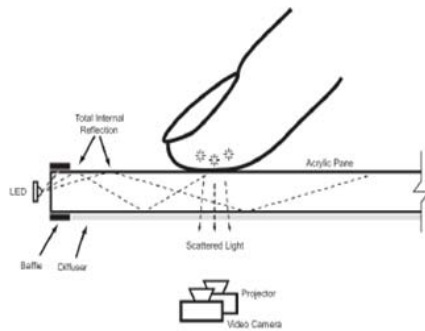


Fig. 2. FTIR: Analysis of the system [14].

construction contains LED that covers the acrylic screen. This light is reflected between the glass and each touch on the surface scatter the LED light down to the ground. Now the system recognizes through the infrared camera under the surface that somebody or something touched the screen and the referring location. After processing the data the projector displays the LCD from below. The advantage of this setup is that the light of the projector cannot be intercepted by the user and a better presentation is possible. For example Microsoft's Surface [1] uses this setup. Another optical tracking technique is Angulation [5]. As the word says, the system calculates the input coordinates with the help of angulations. Cameras are placed on each corner of the tabletop. These cameras record images and those are compared by once again using an image editing software. The distance of the object to each camera is calculated to track it. The program evaluates this information and hands the coordinates of inputs to the system.

3.2 Acoustic Tracking Technique

This tracking technique uses microphones in each corner to record acoustic signals [20]. In addition there are transmitters which send an ultrasound through the surface. Touches or objects on the screen absorb the sound waves and a signal processor in the core of the system calculates the position by using the runtime information and changes of each emitted ultrasound. The advantages of the system are a high resolution because the whole system is covered by the acoustic signal. The high resolution of the screen leads to a high resolution of the touch inputs - more precise inputs are possible. On the other hand the disadvantages are clearly defined: It is simple to interrupt or disturb the system for example through dust or by putting objects on the tabletop [30]. That is why commercial profit remained low.

3.3 Capacitive Tracking Technique

Capacitive tracking is based on a coupling between transmitter and antenna [30]. This kind of system includes a grid of antennas which recognize every change of current. So the change of current caused by the grounding, when for example a finger touches the screen in a certain area tracks the point of input. An example for the capacitive sensor technique is SmartSkin [22] which uses a vertically arranged transmitters and horizontally arranged receivers. As stated above if there is no touching the current remains high but if a touching takes place the current falls because the touching object is grounded and the system recognizes a touch. The embedded antenna grid helps to find

the coordinates of the input. Diamond Touch [10] is another form of capacitive sensor systems. It uses a transmitter grid on the tabletop and receiver on the chair of the user as the result each touch onto the screen completes the circuit and the current is measurable. Furthermore the Diamond-Touch system is able to differentiate between users due activated receiver.

4 PHYSICAL AND DIGITAL REALM

While analyzing computer systems and the ongoing work in the field of human-computer interaction, one has to regard the big difference between physical and digital objects and information. Comparing these two realms is a very difficult thing: At the first look both seem to be very different but if you take a closer look one will find many similarities. Even if the borders between the two realms are fading there are still things either the digital or the physical world has not got or vice versa. The physical realm consists of Trees, Houses, Tables, Computers et cetera. Every object gives back a feedback, like the third axiom of Newton says that every action returns a feedback of equal power (action and reaction) [19]. This is a contrast to the digital world. Here everything consists of ones and zeros. Digital data are relative; it is a certain form of signals or information which are interpreted by the system. Furthermore in the first instance for non-professional computer users digital data are just everything they see on the screen but it is much more. The virtual realm creates nearly an infinite pool of possibilities and the presentation on the screen is just one expression of it. But it is one of the most important ones, human react on excitability and the sense of seeing is one of the most important ones, so the visualization became a huge research area. In contrast to that a visualization of the real world is useless because it gives us already a visual feedback. Following that, seeing both realms and bringing them into consideration one can recognize that researchers try to turn the unreal, digital world into something that is more comparable to the real world [25]. Once the graphical output was created the contest of being more natural started. Today it is sometimes hard to say whether things belong to one or the other realm. Going back to a more physical point of view: The most important difference between both worlds is that physical objects give a tactile feedback as mentioned above and in contrast to that digital objects give back for example visual information. That is the main difference. [28, 7] People who watch photos, for example, can easily sort them, cut them or throw them away. But beyond he or she does not have to start a machine, the person who holds the photo feels them and can easily give them to another person. The tactile feedback of memorabilia is an important factor. Once again regarding the virtual reality on the other hand, we see things that are impossible to create in a physical only world like the playback of moving pictures: Even if we get no tactile feedback it is nice to see a movie in cinemas. It is a middle way that paves the road between both worlds. As well as the hybrid interaction tries to combine advantages and avoid disadvantages.

5 USE OF HYBRID INTERACTIVE SYSTEMS BY DIFFERENT USER GROUPS

With the development of interactive surfaces, hybrid widgets and hybrid interaction the user behavior changed very drastic referring the use of the new systems and computer in general. Regarding users we can categorize them in many different ways for example by time using these systems, by aims and know-how levels or as I split them for this paper: into different working groups.

5.1 Research and Business Users

Research users have a high know-how level and an affinity to those kinds of systems. The use of systems like desktop PCs or notebooks became a standard. But there is one remarkable point: The collaboration with these systems is very tough to realize, interactive work with a group of people is even harder to realize for instance when a single presentation screen and ad hoc alterable data are needed. This opens the way for hybrid systems which combine collaboration between people and the precise of research environment. As Ishii and

Underkoffler stated it in Urp [32] users of this group would immediately use these kinds of system to have a better understanding of what is theoretically talked about. For example in the case of Urp: urban planning and its problems. That was the problem of former interfaces. A graphical presentation on the screen is not comparable to an interactive presentation of digital information linked with physical objects on the surface; a real three-dimensional presentation viewable and alterable by a whole collaboration team. The same you find regarding business users: This user group is locked between a single working station and a multi-user interface for a better team work. Representatives of the work are highly impressed by systems like Urp [32], sources say that they would use so systems for presentation. Other systems like Microsoft Surface [1] support collaboration. These are some reasons the research and business group of users wants to work with interactive systems and hybrid widgets.

5.2 private Users

Generally the know-how level of private users is lower than the level of professionals and the aim of using interactive displays is another: These surfaces serve as social collaboration locations in which users can hang out together e.g. for playing games or exchanging photos. Interactivity systems offer a lot more ways to be creative in a group. An example of former collaboration would be the so called LAN-party where each user had his or her single monitor and data had to be transferred via a network. New systems like Microsoft Surface [1] support multi-user interactions that mean: One single screen displays all data and creates a workspace for a group of people. A wireless data transfer is simple by putting down supported, physical objects, like mobile phones with Bluetooth, on the table [37]. The user becomes more active and single computer usage is replaced by interactive usage in groups. Furthermore the way of input presents a new attraction. Untrained users have many differences using mouse and keyboard for computer control. The change to touch surfaces and the implementation of physical objects opens new ways and lowers the knowledge hurdle. That is why we find many trained and untrained private users using these systems. Because it is something new with an usability and ease-of-use that exceeds the one of standard graphical user interfaces that is embedded on desktop PCs. To state it clearly for this group of users those systems, like the reacTable [17], are much more fun and a lot easier to handle in contrast to graphical user interface controlled systems. As an annotation it is notable that from time to time users of the private group mix up with users of the group stated above.

6 EXAMPLES

In this section I shortly introduce some kinds of hybrid interactive surfaces. All examples show a direct manipulation of digital data by using the screen and widgets as main input device. During my work on this paper I explored some paradigms of interactive surfaces and each one creates a specific aspect of interactivity. The following examples are chronologies:

- Indirect interactions for a direct manipulation of digital data. The digital shadow of physical artifacts creates the spectrum of using possibilities. (e.g. Urp [32])
- Using different physical objects as support for precise interactions on interactive surfaces. (e.g. the PhotoHelix and PhotoLens [7])
- Creating new digital data like music or movies via different inputs through physics and fingers to create a rich diversity of manipulations and to enable a social collaboration system for a playful interaction with the digital realm. (e.g. the reacTable, [17])
- Direct and indirect interactions on screens using physical objects and fingers as input devices to manipulate virtual objects and setting-up a collaboration systems for a multitude of places and people (e.g. Microsoft Surface, [1]).

6.1 Urp: Urban planning tool

Urp [32] is a system that combines interesting tools of urban planning. The most important questions or problems in this field of research are: shadow, proximities, reflection, wind and visual space. Urp tries to connect these problems and finds a solution in form of an interactive tabletop. Some tools the system consists of are the following: a clock object which represents a certain daytime (shadow problem), a wind tool which enables a simulated wind (wind problem), road objects which represents roads or highways near the buildings and a material-transformation tool that for example can change the facade of a building from bricks to glass (reflection problem). (see figure 3). Regarding the wind problem Urp as a weakness: The "generate" wind is just a two-dimensional visual presentation but in realty wind as to be concerned as a three-dimensional influence. With the help of architectural models and a certain tool on the surface a virtual environment is created. The planning tool itself is set-up with the help of an I/O

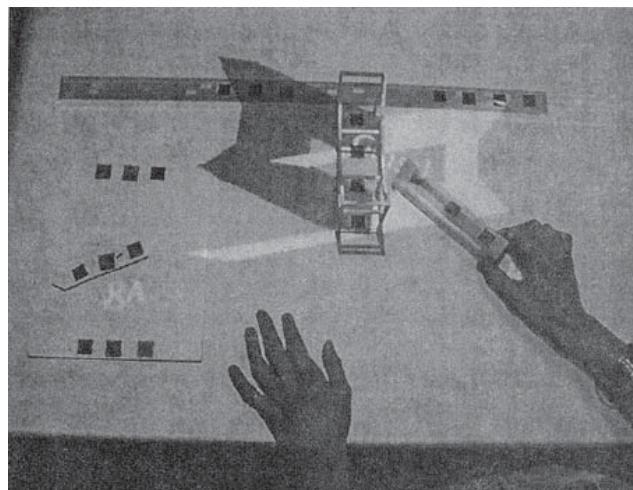


Fig. 3. Changing the digital representation of the facade of from bricks to glass [32].

Bulb [2] and a visual analyzing system (glimpser and voodoo) [31, 4]. The I/O Bulb is a system that projects high resolution visual output and at the same time records a live video of the area of projection. Similar to the construction of Microsoft's Surface [1] the used technology is installed below the surface of the tabletop. The system that evaluates the recorded information is called glimpseer and voodoo as mentioned above. Glimpser on the one hand side tracks colored dots on the surface of each architectural model, voodoo evaluates these information and provides a link between individual points to create the digital form of the model. Urp is a system that combines functionality with visual expression. Professionals who used the planning system were immediately convinced by its possibilities. As Underkoffler and Ishii [32] stated it, they further say that it would be a good learning tool and that it not just might be used in research: Because the displaying of all these useful information would be very attractive for business presentation, too. Non-professional user on the other hand and in addition to that realized a much bigger field of application: for example interior design and the location of objects could be supported by the systems, too. Beyond this all user felt a certain delight using Urp and they came to the conclusion that it has a very low "domain knowledge hurdle" [32].

6.2 PhotoHelix and PhotoLens

PhotoHelix and PhotoLens [7] are both hybrid widgets for the use on interactive surfaces (as shown in figure 4). These two objects were designed for a supporting role while working with images. Both have a supporting role and are used with the non-dominant hand.



Fig. 4. Left: Organizing images into events with PhotoHelix. Right: Scrolling and zooming images with the help of PhotoLens [7].

PhotoHelix

The PhotoHelix is a metallic knob whose upper part can be rotated. It supports image sorting using a timeline and "event folders". The static defined virtual shadow in form of a helix representing the timeline is enabled when the artifact is put onto an interactive surface. In addition to the helix a window is displayed which serves for sorting the images and putting them into consideration with events of the timeline. In fact the position of the widget on the surface is freely selectable. Through turning the upper part the user can scroll through time and he or she is able to select a special event with his or her pen in the dominant hand. Opened an event folder the user can further sort, move or change images.

PhotoLens

In contrast to the PhotoHelix the PhotoLens has a different concept. Although it is a knob too this widget serves for a thumbnail preview of photo piles. Photo piles are a folder metaphor on interactive surfaces. The position of the artifact is freely selectable but the position matters regarding the initialization. When the PhotoLens is put onto a pile of images a window opens with a preview of all containing images. In contrast to that: When it is put onto an image-free area on the surface, the digital shadow will contain a preview folder filled with the images of the clipboard. The left side of the digital shadow will be established right below the widget and it contains a scrollbar which can be used by moving the widget up and down. On the bottom of the window the user can find a clipboard for easy drag and drop operations. In the middle the user can find the preview of all images. Beyond this the PhotoLens has another important feature: When its upper part is turned the preview can be zoomed in or out. This ability supports the functionality of the object and the non-dominant hand. The researchers who implemented these two hybrid widget, Butz et al [7], evaluated the use of both control system as good.

6.3 The reacTable

The reacTable [17] is an interactive table which supports touch inputs and a tangible user interface to provide live music. The project was introduced to the audience at the Audio Engineering Society Conference in Barcelona in 2005. It was designed by members of the Music Technology Group and consists of an input screen and many different pucks. Each can be associated with a unique ability, for example the square shaped ones are audio generators. With putting the different objects onto the table the system creates music by bringing all inputs into consideration (see figure 5). Furthermore the user can manipulate each object by touching its virtual shadow. These shadows contain various settings like speed, frequency or amplitude. So every displayed data is relevant for the music provision. This system uses an infrared camera and a projector beneath the surface for tracking the pucks. The video input of the camera is handled by the reacTIVision - the reacTable vision engine - and sent to the audio synthesizer which creates the audio output and the waveform data for the visual presentation. This reacTIVision as core of the whole system is an open source project and available for the public. Furthermore it is supported by Linux, MacOS and WIN32 [17]. In figure 6 you can see a reacTable set-up. The syntax of the software which the tables uses, is easily to understand: The different shapes are categorized and each one has a special char-



Fig. 5. Bringing different tracked objects into consideration to provide live music [17].

acter and behavior on the music. In addition the system uses a regular grammar (type three of Chomsky hierarchy) and the mixing of global and single affection elements. Regarding these it creates the playful environment of the reacTable as a live audio provision system. Furthermore and like the Microsoft Surface the reacTable creates a new way of collaboration. Users can freely interact whether professional musician or beginner. With the basic knowledge of the input artifacts the user can playfully create live music without guidance. By sliding these physical pucks on the luminous surface and taking different settings for each the interaction of more than only one user creates a new spectrum of interactive music provision. Further studies have shown

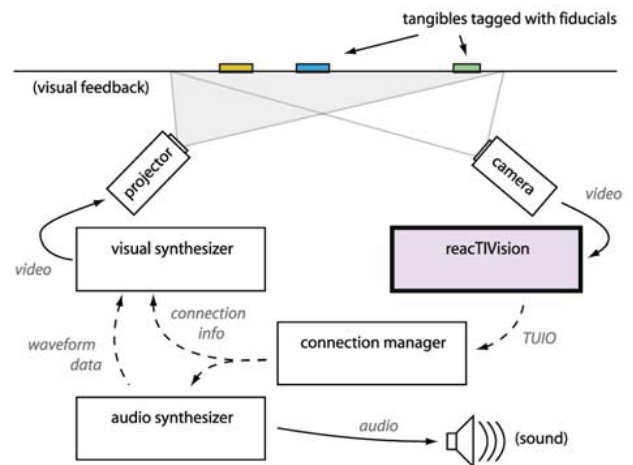


Fig. 6. The construction of the reacTable system [17].

that the reaction of "reacTable musicians" is very positively, the process of creating music becomes a play with an "addiction effect" [17] like other interactive multiplayer games.

6.4 Microsoft Surface

The Microsoft Surface [1] is a table which includes a 30" LCD tabletop that acts as an interactive surface and was first presented in May 2007. The work on the project started in 2001 and the project leaders were Andrew Wilson and Steven Bathiche. The whole technology is set up below the acrylic surface. The user is able to control the product by touching and with the help of several objects. It has a multi-touch and multi-user surface with up to 52 finger inputs at the same time and a detecting system for physical objects like digital cameras, mobile phones or credit cards. Those objects are detected by the system and after a "handshaking process" standard operations are initialized e.g. downloading all photos of the device. The construction consists of an infrared camera, a projector and a computer as you can see in figure 7). The recorded video data, filmed by the infrared camera are evaluated by the computer system inside and the appropriate answers are sent to the projector for displaying them. Furthermore physical objects are

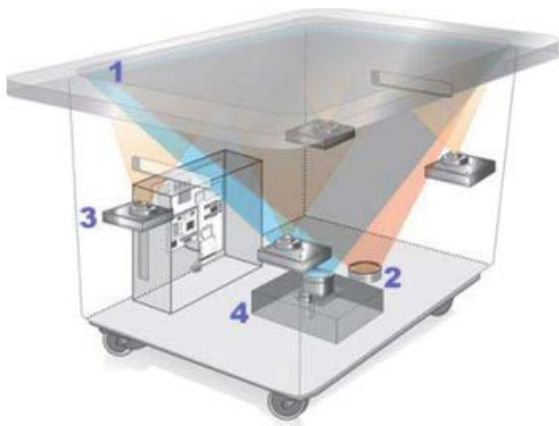


Fig. 7. (1) Diffused acrylic surface as screen, (2) source of infrared light, (3) infrared camera, (4) projector [1].

detected by the camera in the same way. The diversity of supported objects creates the need of differential recognition. [16, 11] For example the system is able to interact with an object that have a special interface like the magnetic strip of a credit card or the wireless LAN function of a digital camera on the one hand side or tracking objects which have a tag on the other. In addition to that, the system includes a lot of interfaces for external, wired, media. The Surface offers a lot of new possibilities. With its dimensions the product provides space for social collaboration. The multi-touch and user interaction support boosts the playful exposure with the apparatus. In addition to the natural inputs the detectable physical artifacts create a new spectrum of interactivity in fields of work and social life. A faster and easier interaction with digital information of different users is possible and their reactions referring to the Surface is widely positive [1].

7 COMPARISON BETWEEN DIFFERENT HYBRID WIDGETS

This part of the paper draws a comparison between the different types of hybrid interactive objects. The presented systems as well as other interactive surfaces which support hybrid interaction, contain an abundant number of detectable physical artifacts. Because of this rich diversity it is advisable to categorize the objects and try to look at them from different points of view. First of all it is needed to create categories for comparing the physicals. In that case one has to regard the most important things of each object. In this work I disregard the form and dimensions of the embedded artifacts; I rather compare the actual way of linking them with digital data as result of hybrid interaction. The link between digital data and a physical object is a question of influence, what changes take place when an artifact is put onto the surface and what is the response of the system? For example some objects do not create a visual response, like a digital shadow, when they are connected. An important question that rises upon the way of linking is the way of manipulation: Does the object need direct or indirect human input and in what aspects does it change the system state? This question is comparable to the first one, it asks for the way in that the inner status of the system is changed, e.g. does an object change the state of the system by directly moving it or does it something on its own without human interaction or interaction on another object. Furthermore the way of tracking objects plays a role for checking supported widgets against others. As mentioned above there are a lot of methods for tracking giving inputs but in this section it is more important to focus at the tracking abilities objects are equipped with. Alongside the above described possibilities, there exist a lot of ways for getting the position of objects and identify their kind. Some of them are: Tagging by using RFID (Radio Frequency Identification) [33] or NFC (Near Field Communication) [24], using dot systems like glimpser and voodoo [31] or different other types of recognition in-

terfaces. After regarding the tracking of objects one have to consider the location on the surface: Does it play an important role where the artifacts actually are put onto or does they do their job no matter what their location is like? The last this paper states for drawing a compare between the objects is the coherence of them on a single system. It is necessary to ask whether the artifacts work as a single unit on the construction or do they need other objects to fulfill their tasks. Finally one can state the difference either the artifact is created for a single system only, so it can be used by only one system or it is simply an object of daily use, for example a mobile phone or credit card, which can be used by more than just one system.

Link between object and digital data

This subsection categorizes the embedded objects by regarding their actual link to digital data. This paper differs between having an own virtual surrounding or just having an indirect influence on the projected information. So we have tabletop systems which support objects with both interactions methods, an example is Urb [32] where the model of the building has no direct influence on the system, just as a tool comes to the surface, a surrounding is created. A counterexample is Microsoft Surface [1]. The meaning of collaboration and ease-of-use let it initialize a virtual shadow for each object which is comparable to the following examples: PhotoHelix and PhotoLens [7] and the reacTable [17]. They try to create a simple usage and a construction of widgets to get a certain result would be circuitous. Single object using systems of the first type with no digital surrounding, would make no sense because hybrid interactivity arises from items which have a certain effect, without these effects the system would be meaningless.

Way of Manipulation

Talking about hybrid widget, another question we have to consider is in what ways the objects are actually linked to the system. Either the widgets generate or can generate an output through moving them or the objects have a passive, static position. Once the current location is defined and a change would affect the whole system. Examples of this group where the moving or rather a fluent interaction of the object by users matters are PhotoHelix and PhotoLens [7] as well as the use of styluses [13]. All these system state it clearly: Objects have to be moved to achieve a certain effect. The objects become direct control elements. ReacTable [17], Surface [1] and some others are hybrid systems regarding this point of view work differently. In those cases it changes from widget to widget whether it plays an active or inactive role. Taking the reacTable as example one can see that first the objects are brought into coherence but once music is provided the system acts without direct manipulation. On the other hand we have this tabletops which just use indirect manipulation methods, like the BlueTable [37] or Urb [32]. Here physics are put onto the surface, the systems reacts and there is no further interaction needed to create a certain output. For example as we use the BlueTable: We put down the mobile phone and if the system recognizes it properly the ongoing interaction is controlled by digital inputs through the surface and no direct control of the artifact is needed.

Tracking

Another way to categorize hybrid system or rather hybrid widgets is the way the objects are tracked: What kinds of techniques are used? Instead of regarding the tracking of the system as describes above, we are now speaking about the tracking techniques of objects. For the sake of brevity I differ between the tagging methods and generalize all other methods to a second group. The standard tagging method consists of a tag which can be doubtless identified by the system. Kinds of these tags vary and there are techniques like RFID [33] or NFC [24] for realizing that. But as soon as such a system is realized another question rises whether tags are truly unique or not. That matters for example at systems which use everyday items because tags of these objects must be unique for a clear identification. There would be the need of a new standard which is always hard to realize. On the other hand tagging for systems which use their own, just for them, created objects is a simple and cheap way of tracking. An example for pure tracking by tags is

Category	Surface [1]	Urp [32]	PhotoHelix & PhotoLens [7]	reactTable [17]	BlueTable [37]
Location	doesn't matter	matters	doesn't matter	matters	doesn't matter
Way of manipulation	direct	indirect	direct	direct	direct
Multiple interactions	no	yes	no	yes	no
Tracking	differs	tagging	tagging	tagging	visual
Link between realms	direct	indirect	direct	direct	direct
Object's life	differs	single	multiple	single	multiple

Table 1. Categories of hybrid widgets

the reactTable [17] where every used object has a single and unique tag. The second group instead uses a various number of different tracking methods. For example Urp [32] uses the glimpse [31] and voodoo [4] technique for the architectural models where each object has different colored dots. Another way is the usage of QR Codes which are a two-dimensional barcode with the help of visual sensor system they can be uniquely identified. This one is for example used by credit cards linked with Microsoft's Surface [1].

Location

As mentioned above the location of widgets can be important for the work of the system. We differentiate between local bound and local unbound. As the name state it local bound means the location of the object plays a certain role and the user have to know that a specific construction of elements is needed to create a certain result or output. That category includes all systems which generate their function out of the position of artifacts. The reactTable [17] is an example which matches this definition. Here all objects are brought into consideration. A variation of the position of one single object can completely change the result, in that case the way of live music [17, 32]. Another example is Urp [32] where the location of the tool is as important as the location of the architectural model. Because both are drawn into coherence to create the output and solving the urban planning questions. A third example is the physical navigation through interactive stories as described by Mazalek et al [18], here the position of the physical objects show the state of the system and they are important for the upcoming events. On the other hand we have systems where the location does not matter. The user can freely put the artifacts wherever he or she wants them to be, as long as they are on the interactive surface the objects can be tracked. The commercially most known example for this type is the Microsoft product called Surface [1]. It support, as described above, a various number of physical object but in contrast to local bounded physics here it does not matter where the user puts the items, e.g. a credit card or digital camera is detected whether it is located in the middle of the table or at one of the corners. Other examples are PhotoHelix and PhotoLens [7] as well as the Bluetable (a system that is able to detect mobile phones and creates a wireless connection) [37].

Multiple Objects

This section has to be split into two different points of view: First we have multiple objects which stand in consideration to each other and we have multiple objects which have no special coherence. Several elements that create one unit are often used at systems which provide a certain output. As well as the examples of the matter of location we find the reactTable [17] and Urp [32] one more time in the same group. Both use multiple objects to generate a solution. Each tool of Urp is considered by the models as well as every puck of the music table is brought into coherence with the whole system to generate live music regarding all types of input (in form of pucks). The second group is once again filled with supporting objects like the ones of Butz et al [7] and Microsoft Surface [1]. Even if the Surface does support a rich diversity of widgets, they do not stand in relation to each other, the credit card does not matter whether the half-empty glass stands next to it or on the other side of the tabletop.

Object types

Aside from form and dimension of several objects the hybrid widgets can be categorized into the group of object which were created just for the purpose of one single system and the group that included the ones which have another function like articles of daily use (mobile phones, digital cameras, etc.). In that case the embedding of the object is a secondary way of utilization because primary they are independent working systems. Systems that support this type of objects are often used as collaboration platforms. They support first of all gadgets which are used by a various number of people and serve as data storage. These are objects like special tagged glasses, credit cards or mobile devices with wireless accesses. An example for supporting the mentioned artifacts and a lot more is Surface [1] or its antecessor Bluetable [37]. In that case objects have their right to exist not just only because of one single system. On the other side systems like Urp [32] or the reactTable [17] where special created tools and pucks take control of the output, the extra for their needs created items have no other destiny than serving this one single system.

Evaluation

This section offers a little round up for the categories mentioned above. Bringing all single categories, as shown in table 1, into consideration we see that it is possible to form groups of physical, hybrid objects.

At first one can recognize a link between whether the location matters or not and if it is about interaction with other objects. The coherence between location and number of artifacts is important for the system and needed for further calculations. Because of the matter between single items, the item itself and their local dependency is characteristic: Is the object locally dependent or locally independent? Furthermore there is a connection between the way of manipulation and the link of the actual physical object with digital data. An indirect way of manipulation generates an implicit link between both realms and vice versa. So another characteristic is established: direct or indirect hybrid widgets. With the help of bringing together the categories to special characteristics we are able to found groups of objects for a better understanding of what they do. First there is the group of local-context objects which stands for location and interaction between objects on a surface and the second group which was generated from way of manipulation and the link between realms includes direct-context elements. These two characterizing groups give us a better understanding of what kinds of hybrid artifacts belong together and how a conclusion of the topic can be drawn.

8 CONCLUSION

The aim of the paper is to give an overview of hybrid interaction on interactive surfaces. First of all and for a better understanding it was important to state the evolution of interactive displays which headed to hybrid forms of interactivity. In addition to the history, a review about the construction and tracking techniques was given to help the reader understand the following sections. The comparison between the virtual and physical realm is the introduction of hybrid interaction. It has shown that both worlds are linked in some kinds of ways but each one has characteristics the other do not have. But for all that the section pictured that the borders with the help of physical artifacts combined with digital information more and more fade. Associated to the construction and the behavior in both realms the importance of user group needs grows. As shown, we can recognize that these systems usually

find use in collaborative working places or as a free-time activity for more than one person. But primarily they serve as support for interactive demonstrations to get a better view of the topic. After regarding several examples we found out that there are a lot of ways to create those kinds of systems and at least there is a big field of opportunities of their usage. It arranges from data transfer over support gadget for precise control to live media provision systems. Because there is such a diversity of objects the last section categorizes them and forms special hybrid widgets groups. As stated in the paper some of the characteristics match and it was easy to form groups of artifacts. So we have the local-context and the direct-context group of utilized artifacts. Objects can be easily wrapped by these groups and this leads to a better understanding of the kind each object consists of. A little interview of students did underline the better understanding of the functionality of the objects with the help of these created groups. By combining the most important facts of hybrid interaction on interactive surfaces the paper attempts to summarize the actual point of research, to give an impression of the difference between digital and physical world and to expand the point of view at hybrid items for a better understanding of hybrid designed systems.

9 ACKNOWLEDGEMENTS

I kindly thank all participants of my short online interview.

REFERENCES

- [1] Microsoft surface. <http://www.microsoft.com/surface>, accessed on 06/01/2009.
- [2] MIT Tangible Media Group. Luminous Room. <http://tangible.media.mit.edu/projects/luminousroom/>, accessed on 06/08/2009.
- [3] Multi-touch technologies. nui group, May 2009. Community Release.
- [4] F. Block, M. Haller, H. Gellersen, C. Gutwin, and M. Billinghurst. Voodosketch - extending interactive surfaces with adaptable interface palettes. *TEI'08*, 2008.
- [5] S. Boring, O. Hilliges, and A. Butz. A wall-sized focus plus context display. *IEEE International Conference on Pervasive Computing and Communications*, 2007.
- [6] S. Brave, H. Ishii, and A. Dahley. Tangible interfaces for remote collaboration and communication. *CSCW'98*, 1998.
- [7] A. Butz, O. Hilliges, L. Terrenghi, and D. Baur. Hybrid widgets on an interactive tabletop.
- [8] W. Buxton, R. Hill, and P. Rowley. Issues and techniques in touch-sensitive tablet input. *ACM'95*, 2007.
- [9] G. Crampton-Smith. The hand that rocks the cradle, 1995.
- [10] P. Dietz and D. Leigh. Diamondtouch: a multi-user touch technology. *ACM'01*, 2001.
- [11] P. H. Dietz and B. D. Eidelson. Surfaceware: Dynamic tagging for microsoft surface. *TEI'09*, 2009.
- [12] F. Ehtler, M. Huber, and G. Klinker. Shadow tracking on multi-touch tables. *AVI'08*, 2008.
- [13] C. Forlines and R. Balakrishnan. Evaluating tactile feedback and direct vs. indirect stylus input in pointing and crossing selection tasks. In *CHI'08*, 2008.
- [14] J. Y. Han. Low-cost multi-touch sensing through frustrated total internal reflection. *ACM'05*, 2005.
- [15] R. Hofer, D. Naeff, and A. Kunz. Flatir: Ftir multi-touch detection on a discrete distributed sensor array. *TEI'09*, 2009.
- [16] S. Izadi, S. Hodges, S. Taylor, D. Rosenfeld, N. Villar, A. Butler, and J. Westhues. Going beyond the display: A surface technology with an electronically switchable diffuser. *UIST'08*, 2008.
- [17] S. Jordá, G. Geiger, A. Alonso, and M. Kaltenbrunner. The reactable: Exploring the synergy between live music performance and tabletop tangible interfaces. *TEI'07*, 2007.
- [18] A. Mazalek, G. Davenport, and H. Ishii. Tangible viewpoints: Physical navigation through interactive stories. *PDC'02*, 2002.
- [19] I. Newton. *Philosophiae Naturalis Principia Mathematica*. E. Halley, 1687.
- [20] J. A. Paradiso. Tracking contact and free gesture across large interactive surfaces, 2003.
- [21] J. Patten and H. Ishii. Mechanical constraints as computational constraints in tabletop tangible interfaces. *CHI'07*, 2007.
- [22] J. Rekimoto. Smartskin: An infrastructure for freehand manipulation on interactive surfaces. *CHI'02*, 2002.
- [23] J. Rekimoto and E. Sciammarella. Toolstone: effective use of the physical manipulation vocabularies of input devices. *UIST'00*, 2000.
- [24] F. Resatsch, S. Karpischeck, U. Sandner, and S. Hamacher. Mobile sales assistant: Nfc for retailers. In *MobileHCI '07: Proceedings of the 9th international conference on Human computer interaction with mobile devices and services*, pages 313–316. ACM, 2007.
- [25] P. H. Rosen.
- [26] D. Schmidt. Design and realization of an interactive multi-touch table. Lancaster University.
- [27] B. Shneiderman. Direct manipulation: A step beyond programming languages. *IEEE Computer*, 16(8):5769, 1983.
- [28] L. Terrenghi. Designing hybrid interactions through an understanding of the affordances of physical and digital technologies. *TEI'07*, 2007.
- [29] L. Terrenghi, H. Richter, D. Kirk, S. Krämer, O. Hilliges, and A. Butz. Physical handles at the interactive surface: Exploring tangibility and its benefits. *AVI'08*, 2008.
- [30] J. Trümper. Multi-touch-systeme und interaktive oberflächen. 2007.
- [31] J. Underkoffler and H. Ishii. Illuminating light: an optical design tool with a luminous-tangible interface. In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 542–549. ACM Press/Addison-Wesley Publishing Co., 1998.
- [32] J. Underkoffler and H. Ishii. Urp: a luminous-tangible workbench for urban planning and design. *CHI'99*, 1999.
- [33] R. Want. The magic of rfid. *Queue*, 2(7):40–48, 2004.
- [34] M. Weiss, J. Wagner, Y. Jansen, R. Jennings, R. Khoshabeh, J. D. Hollan, and J. Borchers. Slap widgets: Bridging the gap between virtual and physical controls on tabletops. *CHI'09*, 2009.
- [35] A. Wexelblat. An approach to natural gesture in virtual environments, 1995.
- [36] A. Wilson and A. Olwal. Surfacefusion: Unobtrusive tracking of everyday objects in tangible user interfaces. *Graphics Interface 2008*, 2008.
- [37] A. Wilson and R. Sarin. Bluetable: Connecting wireless mobile devices on interactive surfaces using vision-based handshaking. *Graphics Interface 2007*, 2007.

Interaction Metaphors on Interactive Surfaces

Renata Willi

Abstract—Over the past years interactive surfaces have emerged as a helpful tool for co-located collaborative tasks. Out of this latest kind of group-working new challenges arise for designers of user interfaces (UI). The aim of this paper is to demonstrate why interaction metaphors are so important for an easy Human-Computer-Communication. It also points out the essential steps of designing new metaphors. Additionally in the main part of this article various examples of interaction metaphors on interactive surfaces are described and illustrated.

Index Terms—interactive surfaces, interaction metaphors, user interface, digital whiteboard, tabletop system, DiamondSpin, paper metaphor

1 INTRODUCTION

Besides the improvements of usual Desktop computers and laptops, interactive surfaces have been developed.

Interactive Surfaces are high-resolution, large-screen, multi-touch displays, which can assume different kinds of shapes and arrangements. The most common examples of interactive displays are vertical digital whiteboards like the *SMARTBoard* or horizontal interactive tabletop systems like the *Microsoft Surface*. Some extraordinary instances are the ball shaped *Microsoft Sphere* or a cylindrical multi-touch interface (see figure 1).

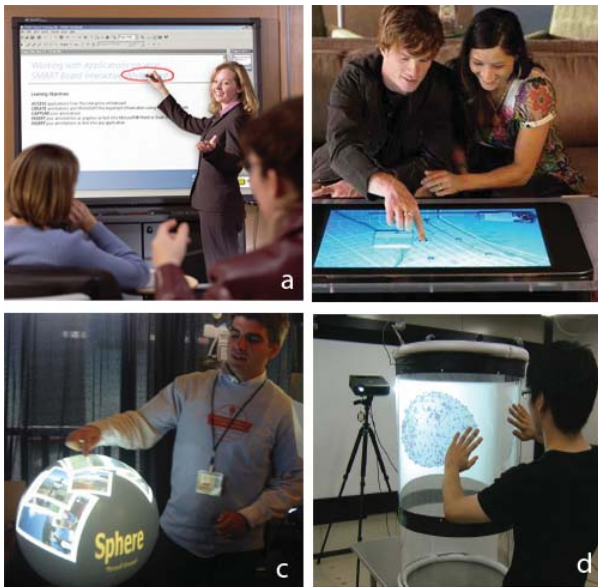


Fig. 1. (a) SmartBoard [20], (b) Microsoft Surface [22], (c) Microsoft Sphere [14], (d) Cylindrical multi-touch interface[19]

Recent developments in display projection have brought the costs of such screens to a level where they can be utilized to support a wide range of everyday activities.

Due to the increased size, large display screens enable a group of people to have a look at the same visual space. As there is only one

screen and the interactive surface is input and output device at the same time the group shares only one single copy of a digital document, which is manipulated directly. So everyone has the same version of it. Consequently interactive surfaces can support both, distance and face-to-face teaching as well as the collaboration, coordination and organization in meetings and discussions. [2]

There are some great differences between common Desktop PCs and interactive surfaces:

- Interactive surfaces are multi-touch facilities. It is also possible that one user has more than one control point at the same time, for example two fingers. In this case it is very important that each user can be identified by the system.v
- Most of the time users interact directly with the system as they would with ordinary whiteboards or papers lying on the table, this means by fingers or digital pens and gestures and not with a keyboard and a mouse.
- "Group work frequently involves transitions between periods of active collaboration and periods of individual activity." [16] Thus the user interface should provide both private and shared workspaces to encourage simultaneous and parallel problem solving.

To meet all these requirements the designers have to "rethink the traditional single-user interface". [6] It is not suitable for multi-touch surfaces. But the real big issue is to let the interaction with a digital user interface disappear and become a part of the human-to-human communication. To face this kind of task UI-designers try to invent new interaction techniques which are specially created for the use on interactive surfaces.

2 USER INTERFACE METAPHORS

The question is what exactly are metaphors and why are they so important? The general term *metaphor* is derived from the Greek word *metaphrein* which means *to transfer*. [7] So you can say metaphors are "mappings from one domain to another".[13]

In computer science metaphors are significant components of the user interface like icons, images, structures or processes. Their goal is to make Human-Computer-Interaction more easy and effective. They are well-designed when the user immediately knows how to interact with the user interface without the help of any tutorials.

To achieve this purpose UI-metaphors work with concepts which are familiar to the user. More precisely this means interaction techniques imitate the look and functions of well-known everyday items. In this way users can transfer their knowledge from their daily lives in order to understand new technologies. This method should reduce novice users' confusion and anxiety towards an unfamiliar environment and increase ease of learning, memorization and usability. "By reducing operational complexity, users may be able to achieve greater

- Renata Willi is studying Media Informatics at the University of Munich, Germany, E-mail: renata.willi@campus.lmu.de
- This research paper was written for the Media Informatics Advanced Seminar on Interactive Surfaces, 2009

initial productivity.” [13] This means better quality of work in less time.

One of the most famous and wide spread interface metaphors is the single-user Desktop Metaphor. It treats the monitor of the computer as if it is the user’s desktop. For instance, documents and folders with documents can be placed, opened, manipulated and closed in the UI-desktop, similar to real life desktops or documents can be deleted by dragging them into the trash can. This can also be compared with the reality of throwing paper into the rubbish bin, unless it is usually not placed on the table.

But as mentioned before this kind of single-user interface does not comply with the requirements of interactive surfaces. New Interaction metaphors should be designed.

3 DESIGN OF NEW INTERACTION METAPHORS

Before creating a new interaction metaphor the designer has to analyze the interactive environment and identify its problems. This means items, structures, processes or navigations which seem to be confusing or inappropriate should be detected. [13]

Such misunderstandings in Human-Computer-Interaction may occur for several reasons:

One is that the metaphor could be quite old, so that the context of the metaphor may have changed over the past years. Consequently it does not fit for its purpose any more and should be replaced. Otherwise the users’ confusion may lead to errors and performance losses.

One example of such an ancient metaphor is the image of the telephone, which has changed remarkably over the time. Back in the beginning and the middle of the 20th century the early telephones had a speaking tube and a circular dial, like it is shown in (picture (a) of figure 2). If you have a look at modern phones, you recognize that the speaking tube has disappeared and the circular dial turned either into push- or touch-screen-buttons, for example on Apple’s iPhone (see picture (b) of figure 2). Today you do not even need any physical device which slightly looks like the ancient phone. One example is the Internet-PC-phone Skype (see picture (c) of figure 2). However, as you see in the (pictures of figure 2) the telephone headset has still survived in many pictographic symbols. Nevertheless the availability of telephones in other forms and devices show the gradual disappearance of its earlier image as an icon reference and the eventual unreadability of the traditional phone metaphor. [13]



Fig. 2. (a) ancient telephone [21], (b) Apple’s iPhone [3] [4], (c) Internet-PC-phone Skype

After this procedure the designer has to find solutions to the problems. He searches for new possible and more appropriate metaphors in the real world. During the creation process the designer always has to have in mind the consumers’ needs.

The Technology Acceptance Model (TAM) presented by Davis [8] describes the important variables which determine the acceptance or rejection of new technologies. Davis says a system will be accepted by the end users, if it satisfies at least two features:

- **Perceived usefulness**, is defined ”as the degree to which a person believes that using a particular system would enhance his or her job performance.” [8]

- **Perceived ease of use**, is defined ”as the degree to which a person believes that using a particular system would be free from effort.” [8]

An example to state Davis’s theory is a staff meeting. The main factor to be considered is time. People would only apply an application, if they believe it helps them perform their job better (Perceived usefulness). Even if the given application is useful, people under pressure would never use a system, if they do not immediately understand how to handle it (Perceived ease of use). Besides the focus in a conference should not be on dealing with the technical system but on finding the solution of the meeting’s discussion.

To create well-designed metaphor some guidelines derived from [11] has to be taken into consideration:

- **”Choose metaphors that provide concrete images”**: Metaphor images should be explained by the icon itself, for instance everyone knows, without a doubt, what a trash can is for.
- **”Metaphors should not rely on mere appearance”**: A concept should not just have the appearance of the real world object, it also should behave like it. This means it should have the same or similar relationships as metaphor like the real world object.
- **”Try to get as close to the original as possible”**: But metaphors are limited, so designers has to add some unique features and properties to the context of the given metaphor which often goes beyond the real-world object’s functionality [13], for example in the desktop metaphor we find menu and task bars which have no counterpart in the real world. Even the trash can is actually not on the desktop.
- **”Metaphors also do not have to be complete, but interfaces need to provide adequate clues to users”**: Leaving out less important concepts or changing them slightly may not significantly affect usability.
- **”Be aware of culture-specific metaphors and concepts”**: Metaphors are used to appeal to aesthetic taste and self-image. At this point it is very crucial to know who the consumers of the metaphor are. Several characteristics like age, gender, profession, national or regional group have to be observed to design the style of the metaphor. Otherwise the new metaphor does not enjoy widespread recognition. For example, you have to develop a metaphor where the customers are business men. You can either choose between a pictographic, a cartoon or realistic style for the metaphor. Actually, the business people would prefer realistic images because they can identify themselves with the metaphor’s style whereas the other two options do not have a serious enough look for them. [13] On the contrary, if you would have to design a metaphor especially for kids, you may have chosen differently.
- **Metaphors not only adequate for one special purpose**: It should be used in various situations. The reason is, the more a person works with a new unfamiliar interface the better he gets to know it. This is especially important for people who have to handle the system even in very high pressure circumstances like decision making in disastrous situations. [1]

At the end of every creation process you have to evaluate the new generated metaphor in a user study in order to get to know what the consumers think about its usability. According to the users’ statements you can improve the metaphor in a second evaluation circle.

4 EXAMPLES OF INTERACTION METAPHORS ON INTERACTIVE SURFACES

In this chapter the theory of creating new interaction metaphors will be illustrated with some practical examples. These metaphors are developed to improve in particular the Human-Computer-Communication on interactive surfaces. The instances are categorized according to what kind of surfaces they are used on, either on horizontal or vertical.

4.1 Horizontal Interactive Surfaces

Although technologies like Skype and other videoconferencing applications are being introduced in users' workplaces, face-to-face communication remains still extremely significant. Horizontal interactive surfaces like tabletops "provide a convenient physical setting for people to meet, chat, look over documents, and carry out tasks that require face-to-face collaboration" [6] as you can see in (figure 3).



Fig. 3. (a) meeting around a real table, (b) meeting around a digital tabletop [18]

First of all we have to define some additional requirements derived from [10] which have to be considered in user interfaces particularly for interactive tabletop systems:

- The meeting members sit around the table facing each other. Thus they have different views on the display as you can see in (picture (b) of figure 3). Therefore it is important that a tabletop UI allows arbitrary document orientation.
- A problem which emerges out of arbitrary document orientation is that the user interface should distinguish between rotatable (e.g. documents) and rotation-sensitive (e.g. menu bars) components
- As mentioned in the beginning of the article, interactive displays are quite large screens. So there are areas and workspace items which can not be reached from every position around the screen. Besides, reaching across the table may lead to territoriality issues during collaboration. This means it is socially awkward to reach across other persons' work or into other persons' workspaces. Therefore tools for document duplication, sharing and passing are necessary.
- Another unique feature of tables, which have to be taken into account, is that they do not have a fixed size or shape. They even do not always have a fixed number of persons sitting around it on fixed positions. As others join and leave the group people may move round the table and adjust their seating arrangements. For this purpose it is very important that the private and public workspaces are adjustable within the user interface.
- The users need to work with the documents on the table and manipulate them. This usually happens inside their personal workspaces. Actually these personal workspaces on their own, just need a single-user interface, because only one person works in it. Here the common Desktop metaphor can be used with its manipulation tools.

To satisfy all these demands the user interface needs many different kinds of tools and applications. Not to disturb the users' communication the designers should pay attention that the switching from one activity to another would be fast and easy. [17]

4.1.1 DiamondSpin Application

The DiamondSpin Toolkit with its own DiamondSpin API is not an interaction metaphor itself, but it is often used to implement multi-user

interface prototypes for tabletop systems very quickly. At the moment DiamondSpin is freely licensed for academic researchers to enable more people to develop and explore new interaction metaphors and so boost the improvements in UIs for interactive surfaces. Therefore to be platform independent the DiamondSpin framework is written in Java/Swing.

This toolkit solves the problem of arbitrary viewing angles, because it is based on a Polar- and not on a Cartesian- coordinate system, which has only one common direction of orientation for its visual objects. This kind of coordinate system enables an easy rotation of visual components as well as the rotation of the entire tabletop surface.

One problem goes hand in hand with the rotation of items. The system has to make a difference between rotatable and rotation-sensitive parts of the screen. The solution is a multi-layer architecture. The lowest layer contains non-interactive components like menu bars or maps. The topmost level controls the few active components which are currently receiving user events like passing or rotating a document.

This kind of architecture simultaneously solves another multi-user interface problem - the performance and response time on users' input. As all users interact with the surface at the same time, the system may have some performance issues while re-freshing the screen. Most of the time only currently active components have to be updated. This means just the topmost layer has to be repainted frequently, while the other layers remain the same.

DiamondSpin already implements various important features to face tabletop challenges such as color coded frames to distinguish between different window owners, assumed the tabletop system provides user identification information. It also supplies relocatable and non-modal menu bars for adjusting the seating arrangements of users and various sizes and shapes of different tabletops. The rest of the table remains public area. At least it also supports free-hand annotation mechanisms. [18]

The following four interaction techniques [16] were prototyped using DiamondSpin:

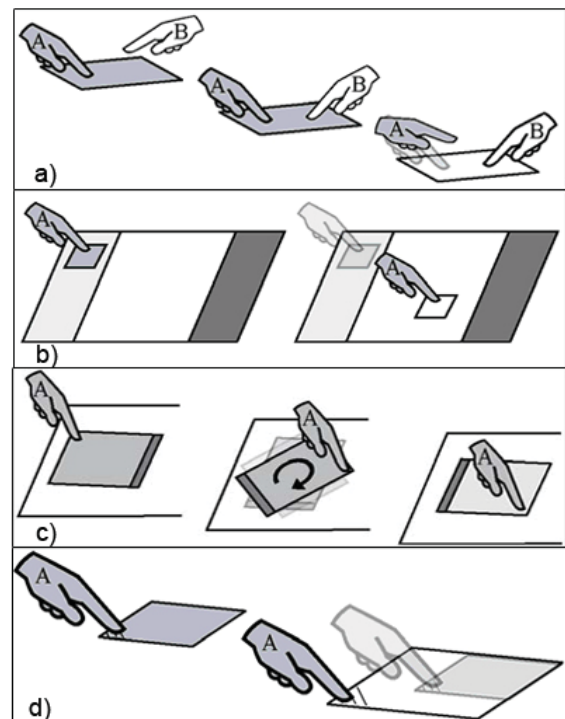


Fig. 4. (a) Release, (b) Relocate, (c) Reorient, (d) Resize [16]

- **Release:** In (picture (a) of figure 4) you see that user A holds an electronic document. User B wants to take it away. He only will

get the document, if user A releases it, otherwise it will remain in users' A property.

- **Relocate:** As there are personal and public areas on the table documents can be associated to each region. If a document is moved into a public area it will turn to public mode, while moving it to a user-owned region makes it private. You can not only relocate surface components but also assign and relocate private areas. By touching the tabletop surface, the area closest to the user appears as his or her private workspace. When leaving the table the user can double-tap his or her private region to make it public again. This procedure is illustrated in (picture (b) of figure 4)
- **Reorient:** This tool allows changing the orientation of digital documents like real-world documents lying on the table. The orientation of the document defines its mode. If a document is oriented towards the center of the table, it is public. But if it is looks towards the outside to the user who owns it, it is transferred to personal mode. (picture (c) of figure 4)
- **Resize:** With this technique a user can gain a document. Reducing a document's size smaller than a threshold, makes it private, "while enlarging it opens it to public access." [16] All of these four tools secure that only one person is able to manipulate a document. Therefore conflicts between different versions cannot occur. (picture (d) of figure 4)

The following two applications are also implemented with the aid of DiamondSpin.

Table For N - A small group of people (two to four) are sitting around a digital tabletop system and sketching, manipulating and browsing various types of documents, including text, html, images and video clips, with Table For N. But this application only provides maximum five different polygonal views on the display (four private workspaces and one shared workspace in the center of the screen which is illustrated in (picture (3) of figure 5)). Every person can use this tool to achieve a perfect orientated view on his current document; workspace rotations of a value of 45, 90, 180 and 270 are also possible. [18]

Poetry Table - as you can see in (picture (2) of figure 5) is an educational game which allows up to four simultaneous players to combine a set of English or Japanese word tiles to create a poet. These tiles rotate automatically to face each side of a rectangular tabletop. If there is a huge request of a popular tile it can be easily duplicated by pop-up menus supported by DiamondSpin. It is also possible to make a screenshot of the game to save the poems. [18]

4.1.2 Paper Metaphor

People in group meetings typically use pencils and papers for activities like brainstorming or discussing. In order to facilitate the transition to tabletop systems, the same intuitiveness as using paper should be offered for using digital documents. To achieve this aim characteristics of paper like folding can be used to create new interaction techniques.

The following two metaphors are based on this paper metaphor. They are called peeling and slots. As mentioned before, the DiamondSpin toolkit is used to quickly create user interface prototypes. Peeling and Slots are also one of those DiamondSpin based metaphors. [6]

4.1.3 Peel Metaphor

The peeling metaphor allows the user to fold a digital document like a real piece of paper. While this document is peeled the user can see the other UI-components lying underneath. This method can be used to re-organize the users' workspace and to make room for other items. The folded state is only temporary, while the user needs the extra space for working.

As this interaction technique is new for the consumer, the handling of it needs to be obvious and natural. On the one hand this is important because this method can be applied to almost every type of document. On the other hand it is not very functional to hide its activation by

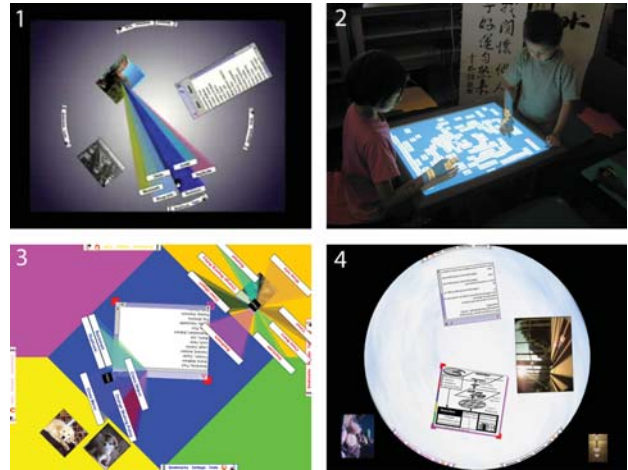


Fig. 5. (1) Rectangular tabletop with continuous document orientation, (2) Poetry Table, (3) tabletop with four private and one public workspace in the center, (4) rotatable circular tabletop [18]

some complicated gesture. According to the Technology Acceptance Model the users will not search for it, when they do not immediately understand the usage and handling of the new technology.

There are three ways proposed to activate the peeling, which are demonstrated in the three pictures of (figure 6):

1. The user can drag one edge of the document in order to fold it in the direction where he makes a movement. The other corner has to be static. On the surface the static corner is illustrated by a pin, the four arrows show as usual the location of the moving point and the grey lines indicate the user's movement (see left picture of figure 6).
2. You can not only peel the edge of a document but also its borders. As you can see the document can be peeled in any shape you want. You can also peel it more than once (look at the middle picture of figure 6)
3. The third interaction mode happens automatically to avoid collision. When a document hits an obstacle such as another item or the border of the table, the document peels itself like in the (right picture of figure 6)

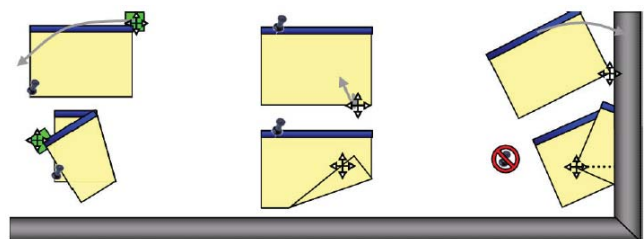


Fig. 6. The three methods of peeling a document [6]

Going a little bit further you can use this peel metaphor as well for grouping and classifying numerous documents. As a result the pile can be moved, rotated, zoomed or easily exchanged with others as a whole like a single document.

In real world such piles surrounded by a folded piece of paper can be found on many desks similar to the one in (figure 7). This informal way of piling documents is complementary to the folder system. A pile can assume two different states: one when a pile is created and the other one, when the documents of the resulting pile are manipulated.

In the first phase there are two techniques to make a peeled pile: Either you fold the pile handler first and put some documents into it or you make a pile of documents and put a pile handler around it. As you see, it works like in the real world. It is important to give the user a clear feedback that tells him which document is about to be piled. But it is also vital for the user to know which pile is about to receive a new document. This information is then essential for the user, if he actually does not want to pile the currently holded document or changes his mind before piling.

In contrary, we need to break with the metaphor in the second phase. As we work with digital data we can use the information, documents carry such as metadata and document type. According to this information tasks can be performed on the piles, for instance a stack of images could launch a slideshow or a pile of graded homework can calculate the mean value of all grades and display them in a chart.

The metaphor can be dissolved by moving a piece or a group of documents out of the pile handler or by removing the folded handler. [6] [9]



Fig. 7. Physically and digitally piled documents [9]

4.1.4 Slot Metaphor

The slot metaphor can be compared to a lot of different kinds of slots in the real world, for instance a letterbox, a copy machine or a paper shredder. These metaphors are easy to understand because they have the same effect in both the digital and the real world.

For example, the letterbox slot provides a good opportunity to meet the need of spreading documents. Elements can not only be transmitted from one side of a table to the other but also between two different tables by addressing the right person on the right tabletop. This also solves the problem of moving or getting elements to or from an unreachable area of an interactive surface. The copy slot duplicates documents and the shredder slot deletes elements.

The interaction with the slot metaphor is quite intuitive (see figure 8). You have to move a document over the gap and activate the slot by pushing a button. Otherwise it would be possible to pull a document in a slot even if the user does not want to use this slot. He just moved the document accidentally over the slot. Not only sending but also receiving a document happens on demand. The special feature when getting a document from someone is that the item is flipped during transmission and is correctly orientated on the receiver's side.

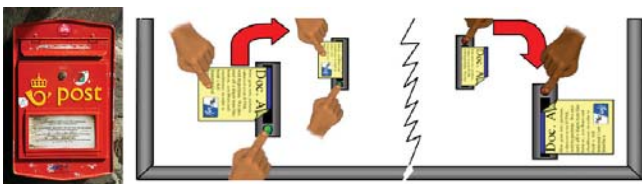


Fig. 8. Sending and receiving a document with a letterbox slot [9]

But there are also not as obvious slots as the letterbox. Besides, free-hand annotation mechanisms can be improved by the slot metaphor. It could be used either to erase annotations or change their color or recall old ones or reveal colleague's annotations. [9]

4.1.5 Interface Currents

Interface Currents are a fluid interaction technique and can be seen as flexible containers for surface items. The elements inside containers are moved by a controllable flow along a path inside the area of the container. This enables intuitive organization and sharing of digital data. Real world metaphors for Interface Currents are for example (see figure 9):

- Conveyer belts in Sushi restaurants to distribute food to customers seated at various locations next to the Sushi bar
- Luggage carousels at the airport to distribute the luggage among the passengers.
- Lazy Susans which are often found in Asian restaurants to share food at a table but also applied to kitchen cupboards to have a bigger storage room.



Fig. 9. from left to right: conveyer belt at a sushi bar, luggage carousel at the airport, Lazy Susan in an Asian restaurant and a carousel in a kitchen cupboard [10]

As the name implies, Interface Currents look like fluid currents with a continuous onward movement. Therefore they can be determined by its flow and path. The flow has a specific direction and speed. Both values can be adjusted by the user by interaction with the interface. A current flows along a path and is surrounded by some boundaries, which defines the currents shape and size. These two variables can also be controlled by the user. Workspace items such as documents and images can be placed on the Current and be affected by its flow.

Currents can be re-located everywhere on the screen. They can even be positioned so that they are only partly visible on the display. The flow can be used for scrolling so that all items can be seen some time. This kind of current is comparable with a carousel in kitchen cupboards. There are two main types of Currents. On the one hand the so called pools and on the other hand the streams. You can see both of them in (figure 10) They are not only different in their appearance, they also support various types of interactions and tasks. **Pools** are like Lazy Susans and have just one exterior boundary. By default the border of the pool is a circle. But it is adjustable to any shape by touching and moving its boundary control points. A pool can also be extended or shrunk as well as arbitrarily re-located. In pools items are oriented so that the elements inside are always looking to the outside boundary. Once the pool is rotating the items follow the outer border, as well as their orientation. Pools provide either areas for the group to share their documents or for personal storage spaces. Storage spaces hold many personal items to provide a clean workspace. Streams are similar like rivers or conveyer belts. In contrary to pools, streams have both, an exterior and an interior border, which can be again adjusted with the aid of the boundary control points. Items move parallel to its boundaries. The elements' size is adjusted to the width of the current at every place (look at picture (a - c) of figure 11). Streams can be stretched out over the entire tabletop surface. So the items flow nearby all persons, so that everybody at the table has good access to them.

Interaction with Interface Currents:

- Manipulating the currents' borders by its boundary control points (picture (a + b) of figure 11)
- Manipulation of the current's position. A current can be moved to any location in the workspace even partially outside the surface by simply touching and dragging its outer border. If an Interface Current is used as personal storage area, it can be reasonable to move it partly out of the working area in order to save

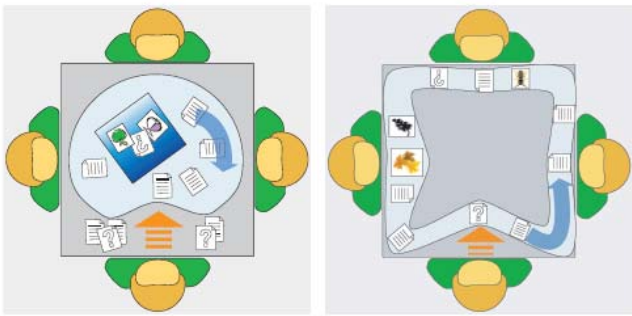


Fig. 10. from left to right: conveyer belt at a sushi bar, luggage carousel at the airport, Lazy Susan in an Asian restaurant and a carousel in a kitchen cupboard [10]

some space. But if it is used for collaborative storage, it should be positioned either as pool in the center of the table or as current which flows around the table so that everybody has perfect access to the items. (picture (d - f) of figure 11)

- Adjusting the flow of a current. You can change the direction of the flow by touching inside the boundaries and giving the flow a new direction. The length of your movement will determine the speed of the flow. With a brief touch inside the Current the user can stop it. With another brief contact the flow starts again.
- More than one Interface Current in the workspace. These currents may overlap sometimes, but the items of both currents would not mix up.
- A user can put into and remove elements from an interface current. If the Current is empty, the flow is not visible.
- By double-clicking on a folder, it is transferred to a current. This enables users to share the content of a folder more easily among each other.

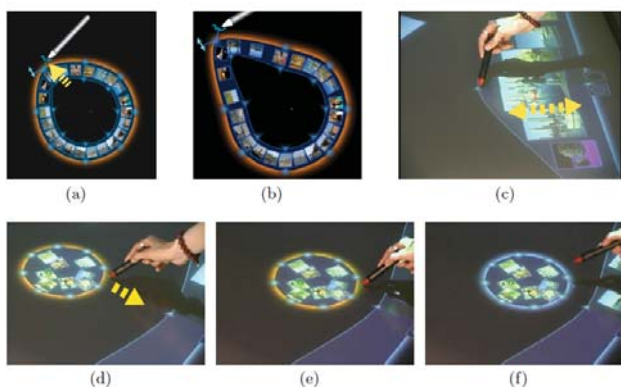


Fig. 11. Interaction with Currents: (a + b) adjusting a stream's shape both inner and outer boundary, (c) adjusting only the inner boundary of a stream, (d - f) relocating a Current (pool) [10]

All in all Interface Currents facilitate through the combination of motion on demand and high degree of flexibility, information sharing and access on large interactive tabletop surfaces. It also supports the teams' mobility in a very easy way.[10]

4.2 Vertical Interactive Surfaces

Vertical interactive surfaces have a stronger presentation character than horizontal ones. They enable the presenter to share one visual focus point with the audience. Therefore these kinds of screens can easily be used to support teaching, both face-to-face and distance learning, as well as business presentations. Most vertical interactive surfaces use the whiteboard or blackboard metaphor as you will see in the next paragraphs. [2]

4.2.1 eChalk / Chaklets

eChalk is a user interface for digital wall displays. It is based on the old-fashioned chalkboard metaphor. Like a common chalkboard it can be used in lessons. As the teacher writes the ideas of the lecture on the board by hand, the pace of the tutorial is much slower compared to presenting materials on prepared slides. This is an advantage for the students, because they are not overwhelmed with too rapidly presented information. This often results in a deeper understanding. Nevertheless eChalk is an electronic device. Therefore it has all possibilities to display multimedia-objects like video streams to enrich the lessons. Another great convenience is that all actions on the board can be tracked easily. The recording is not only limited to the screen, but can also capture the teacher's voice. This is the reason why this kind of technology can be perfectly used for distance teaching as the record can be watched both as a live stream and as a replay. JavaApplets can also be shown on the board to demonstrate difficult processes for example in mathematics. But it has to be taken into consideration that the whole lesson is transferred over the internet, where not every student has access to a high bandwidth transmission, which the presentation of complex JavaApplet require. Therefore a new kind of application has to be invented, which is based exclusively on board strokes - Chalklets. "They are, as the name implies, very similar to Applets, in the sense that they are small Java applications dynamically embedded in a host application." [12] The Chalklet Framework can be compared with the DiamondSpin Toolkit. Developers who are interested to create a new Chalklet application can download the framework from website of the free university of Berlin [5].

To demonstrate what kind of possible application can be implemented, here are some examples:

- **Python interpreter Chalklet** - algorithms and programming can be taught interactively on an eChalk board. The Python Chalklet accepts handwritten Python commands which are recognized and interpreted by the Chalklet as you can see in (figure 12)
- **Graphical simulator for logic circuits** - the user draws the circuit onto the board. Every different items have various colors, which gives a visual feedback to the user. The application recognizes the different symbols for the logic gates *and*, *or*, *not*, *multiplexer* and *demultiplexer*. For simulating the sketched logic circuit, there are two possibilities. First, to start the simulation by drawing a one (high) or a zero (low) near a logic level input. The system sets the nearest input of the circuit to the corresponding state. Alternatively the user can also push the Button "RUN" to start the simulation. (figure 12)

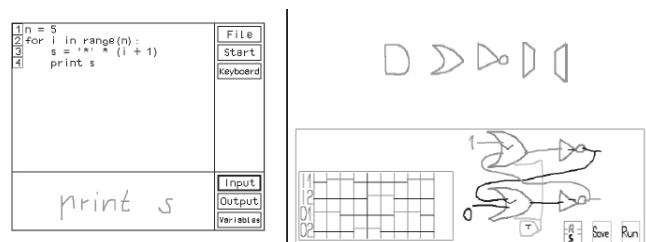


Fig. 12. Two Chalklet applications: left: Python interpreter, right: Graphical simulator for logic circuits [5]

As you see in both Chalklet pictures the chalkboard look-and-feel was retained. On the one hand to keep up with the old-fashioned chalkboard and on the other hand low-level graphic objects do not need as much bandwidth as high-resolution illustrations. [5]

4.2.2 Tivoli Application

Tivoli is an electronic whiteboard application which runs on the pen-based Xerox LiveBoard. As it simulates whiteboard functionality, it naturally offers scribbling and erasing interactions, but it is also possible to edit material by gesturing, which are interpreted pen strokes. The specialty with Tivoli and its gestures is that scribbling and gesturing are separate modes, so that any kind of ink stroke can be drawn without any predefined gesture strokes.

Unlike a whiteboard it also provides interaction techniques for organizing and rearranging materials as well as for managing the board space. They are quite useful, because ideas do not arise as a whole at once. It is usually worked out step by step. Another important feature is the UNDO-functionality. Users can write on the board without any fear, knowing that they can back up, if they have to. By this Tivoli encourages informal and fluid communication in group meetings.

Tivoli has a single surface approach. This means the board has just one surface. Thus objects cannot be hidden by overlapping surfaces.

As the digital whiteboard is an electronic device, there are unlimited scrollable pages. But it is more important how much material (ink strokes, gesture strokes, characters and icons) is visible at the same time. To use the board space efficiently the users can group materials into categories and treat them as a unit. Materials are grouped together by creating and manipulating boundaries and divide the visual surface into distinct regions. As you see, Tivoli techniques are based on "direct manipulation of boundaries and the implicit recognition of regions." [15]

The following paragraphs describe several Tivoli interaction techniques for the organization of materials on the board:

- **Basic selection and editing** - Selecting objects on the board is easy and intuitive to handle. The user has to surround the desired objects with a gesture stroke to select them. These selection loops are only temporary boundaries, which can be altered. To reshape a selection, the gesture must begin and end on the loop. Between these two contact points the loop is modified either to the outside, called bump, or to the inside, called bite (figure 13). Tivoli facilitates the basic editing operations like move (drag gesture), copy (gesture like a "C") and delete (pigtail gesture). To make some room on the board scaling operations are advantageous.



Fig. 13. "A freeform selection loop is altered by a bump gesture to include new material" [15]

- **Structured borders** - As mentioned above, the organization of the board is managed by arbitrary borders as you see in (figure 14) The user can generate a border with a horizontal- or a vertical-line gesture. Borders can be manipulated with the same gestures as materials. The tap-gesture selects a border, the

pigtail-gesture deletes borders and the drag gesture moves borders, but only until it hits an object. This prevents a user to move objects accidentally to different regions. But if objects are moved inside a region, the borders are adjusted to the new space the objects need. Similar to borders are gaps, which proved to be useful to insert items into already written lists. A gap is also created by a horizontal- or vertical-line like a border (look at figure 15)). The question is how the system can differentiate between a gap and a border. A border is only drawn when there is already a gap.

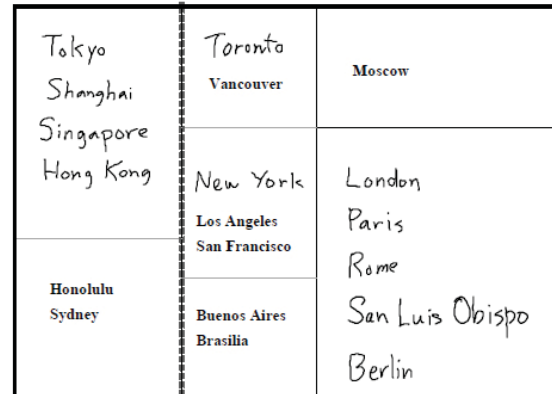


Fig. 14. Structured border is selected. Cannot be moved beyond the "g" in "Hong" or the "N" in "NeyYork" [15]

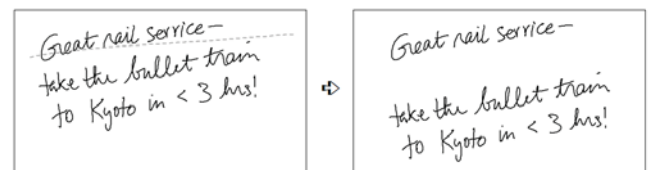


Fig. 15. Horizontal-line gesture creates a horizontal gap [15]

- **Freeform enclosures** - These structures are to group objects really fast, when the user does not want to tessellate the whole screen. An enclosure is created by just drawing a stroke. Thus it is like a graphical object, but it groups the inside objects to a region. An enclosure can either be selected by a selection loop, as described before, or just by a tap on its stroke. Afterwards it can be dragged with its content along the board. Of course, enclosures can be altered. This works just as the gestures altering selection loops. But enclosures can also be splitted (back-and-forth gesture) or fused (loop gesture which cuts through the desired enclosures). Enclosures can also have relationships among each other.
- **Collapsible annotations** - Users often make annotations to keep something in mind. But annotations take lots of space from the board, although it is not important to see them all the time. Therefore annotations in Tivoli are collapsible. They can be minimized into a footnote by a balloon gesture. A Double-tap on the footnote displays its content in an overlay. These procedure is shown in (look at figure 16))

5 CONCLUSION

In this paper I explained why metaphors are so vital for a good user interface. As the technology of interactive surfaces is quite young, it is not seen quite often in daily life. The usability of such interactive surfaces is still not good enough to use them for serious tasks. By now, those kinds of surfaces are merely used to look at pictures. But

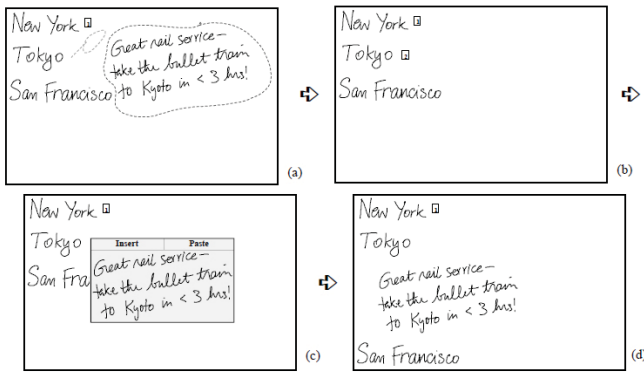


Fig. 16. (a) Ballon-gesture shrinks the selection into a footnote, (b + c) Double-tap on the footnote displays its content in an overlay, (d) "Insert" inserts the contents into the list [15]

there are lots of different attempts to create and establish new interaction metaphors for interactive surfaces to change the status of this technology. However, it will take its time, as there are hardly any user evaluations, yet. But with toolkits like DiamondSpin and Chaklet researchers are on the right way, because these frameworks enable fast prototyping and evaluation of new ideas and thereby reduce the developmental time.

REFERENCES

- [1] H. S. S. Z. Alessandra Scotta, Harmen Hofstra. Multi-user tangible interfaces for effective decision-making in disaster management. In *Remote Sensing and GIS Technologies for Monitoring and Prediction of Disasters*, pages 243–266, Berlin, Berlin, Germany, 2008. Springer Berlin Heidelberg.
- [2] M. Apperley and M. Masoodian. Supporting collaboration and engagement using a whiteboard-like display. Technical report, Department of Computer Science, University of Waikato, Hamilton, New Zealand, 2000.
- [3] Apple. iphone. <http://www.flickr.com/photos/nielsvaneck/1490570992/>, 2009. visited 15.07.2009.
- [4] Apple. iphone. <http://www.flickr.com/photos/powerbooktrance/353462860/sizes/o/>, 2009. visited 15.07.2009.
- [5] F. U. Berlin. E-chalk. <http://kazan.inf.fu-berlin.de/echalk/de/chalkletTutorial.php>, 2009. visited 13.07.2009.
- [6] G. Besacier, G. Rey, M. Najm, S. Buisine, and F. Vernier. Paper metaphor for tabletop interaction design. In J. A. Jacko, editor, *HCI (2)*, volume 4551 of *Lecture Notes in Computer Science*, pages 758–767. Springer, 2007.
- [7] Brockhaus. Definition: Metapher. <http://www.brockhaus.de/suche/index.php?begriff=metapher&bereich=mixed&x=0&y=0>, 2009. visited 13.07.2009.
- [8] F. Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. In *Perceived usefulness, perceived ease of use, and user acceptance of information technology*. MIS Quarterly, 1989.
- [9] F. V. Guillaume Besacier. Peeling for piling: Casual organization of documents. 2007.
- [10] U. Hinrichs, S. Carpendale, and S. D. Scott. Interface currents: supporting fluent face-to-face collaboration. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Sketches*, page 142, New York, NY, USA, 2005. ACM.
- [11] W. Hudson. The whiteboard: metaphor: a double-edged sword. *interactions*, 7(3):11–15, 2000.
- [12] L. Knipping and M. Liwicki. Chalklets: Developing applications for a board environment. In *ISM '06: Proceedings of the Eighth IEEE International Symposium on Multimedia*, pages 907–914, Washington, DC, USA, 2006. IEEE Computer Society.
- [13] A. Marcus. Metaphor design in user interfaces. *SIGDOC Asterisk J. Comput. Doc.*, 22(2):43–57, 1998.
- [14] Microsoft. Microsoft sphere. <http://research.microsoft.com>, 2009. visited 13.07.2009.
- [15] T. P. Moran, P. Chiu, and W. van Melle. Pen-based interaction techniques for organizing material on an electronic whiteboard. In *UIST '97: Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 45–54, New York, NY, USA, 1997. ACM.

- [16] M. Ringel, K. Ryall, C. Shen, C. Forlines, and F. Vernier. Release, relocate, reorient, resize: fluid techniques for document sharing on multi-user interactive tables. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 1441–1444, New York, NY, USA, 2004. ACM.
- [17] S. D. Scott, K. D. Grant, and R. L. Mandryk. System guidelines for co-located, collaborative work on a tabletop display. In *ECSCW'03: Proceedings of the eighth conference on European Conference on Computer Supported Cooperative Work*, pages 159–178, Norwell, MA, USA, 2003. Kluwer Academic Publishers.
- [18] C. Shen, F. D. Vernier, C. Forlines, and M. Ringel. Diamondspin: an extensible toolkit for around-the-table interaction. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 167–174, New York, NY, USA, 2004. ACM.
- [19] B. Shizuki, M. Naito, and J. Tanaka. Browsing 3d media using cylindrical multi-touch interface. In *ISM '08: Proceedings of the 2008 Tenth IEEE International Symposium on Multimedia*, pages 489–490, Washington, DC, USA, 2008. IEEE Computer Society.
- [20] SmartBoard. Smartboard. www.smartboard.de, 2009. visited 13.07.2009.
- [21] stockxpert. phone. http://www.stockxpert.com/browse_image/view/43430611/?ref=sxc_hu, 2009. visited 15.07.2009.
- [22] techshout. Microsoft sphere. <http://www.techshout.com>, 2009. visited 13.07.2009.

Interactive Displays for Creativity Support

Felix Lauber

Abstract— This paper presents an overview of research projects in the application field of creativity supporting environments using interactive displays. Interactive displays are very suitable for collaborative working processes. This is the reason why the main emphasis of both, this paper and the research projects referred to, lays on creative processes executed within a group of users. The design of systems which support creativity is very challenging. This applies especially to systems that support collaboration at the same time, as group working processes can be very dynamic and complex. Designing such a system means to consider the support of creativity itself as well as the assistance of group related processes like conversation and discussions among the participants, the spatial dispersion of the users or different phases in group work.

This paper tries to enhance the comprehension of the problems system designers are confronted with. Therefore we present five attributes of applications supporting creativity by using interactive displays. Each attribute is introduced by giving two representative implementations in research projects in order to provide a possible solution at the same time. In detail these attributes are the *separation of shared spaces and private spaces*, the *fast, direct and concurrent interaction by all contributors at any time*, the *direct communication link of participants among each other*, the *physicality of objects* and the *automated support of browsing and structuring information*.

Index Terms—Creativity, collaboration, interactive environments, displays, overview

1 INTRODUCTION

Creativity supporting environments can be used in a wide range of application scenarios, for example product development, business meetings where ideas are presented and discussed with other colleagues and all occasions where collaborative problem solving can enhance the quality of work.

The design and development of user interfaces, applications and supporting hardware is a challenging task and there is much ongoing research in this domain. In order to improve traditional creative processes there is the attempt to integrate digital technology in non-digital working processes already established, like for example brainstorming. Hilliges et al. stated that the goal should be the joining of the advantages of digital technology - like for example persistent data storage, easy information access and the possibility to navigate backwards in the working flow - and those of interpersonal communication and interaction[4].

Shneiderman separates a creative process into four phases: *Collect, relate, create* and *donate*[12]. In the first phase the user tries to find information related to his task, for example in the web, in databases or in libraries. In the second phase he consults peers and mentors to discuss the topic and get even more information on it. The create-phase is the phase where the user is really creative in the common sense: Shneiderman suggests actions like *thinking by free associations*, *exploring solutions-what-if tools*, *composing artifacts and performances* and *reviewing and replaying session histories* assigned to this phase. In the last phase, the donate-phase, the results of work are disseminated and presented to others.

The focus of this paper are applications supporting creative processes by using interactive displays. Interactive displays seem to be very suitable not only for creativity support but also to support collaboration activities among several users. Reasons for this are probably the possibility of visualizing information in a way that many people can attend, the intuitive interaction possibilities and that they are not constraining the users from moving and communicating freely. That is why most of the applications presented in this paper not only include collaboration modes but are mostly based on the idea of being creative within a group of people. A very good example of an applica-

tion supporting several typical collaborative working scenarios, is the iLAND-project[13]. This project interprets the idea of collaboration in a very comprehensive sense: collaboration is not only possible with many people working on one device but also the different system components are designed in a way that the full system capabilities would not be fully developed until the components are connected with each other and used together.

In this context we tried to elaborate five typical attributes which are considered as being necessary or at least extremely useful for applications supporting those processes in a way which increases the productivity of each participant. This can happen in augmenting the quantity and quality of the achievements of each participant but also in improving special group related processes like communication and discussion in general or the selective support of several phases of group work. While the first three attributes are considered as being useful for collaborative processes within a group of people in general, the last two are supporting creative processes technologically. These attributes are in detail the *separation of shared spaces and private spaces*, the *fast, direct and concurrent interaction by all contributors at any time*, the *direct communication link of participants among each other*, the *physicality of objects* and the *automated support of browsing and structuring information*.

Each attribute is introduced by presenting two applications which implement the attribute and thus can help to explain it.

2 SHARED SPACES AND PRIVATE SPACES

During collaborative creative working processes users enter different working phases. Hilliges et al. differentiate between a generative phase, where users create their ideas on their own and a structural phase, where ideas are presented to the other members of the working group, being discussed and structured[4]. Consequently developers of creativity supporting applications try to facilitate the performance of these phases directly within the application. Introducing dedicated spaces, one for each of the phases, is an approach one can find in many creativity supporting applications. This applies especially to applications using interactive displays, as users have different requirements to the display according to the active working phase: in the generative phase they want to work alone or in a small subset of the group, while in the structural phase all group members need to have access to the display.

While in the brainstorming system of Hilliges et al. the two phases were spatially separated by having two working devices (the table and the vertical display wall), another interesting approach to support this fact is given in RoomPlanner[16]. This project has its emphasis on

-
- Felix Lauber is studying Media Informatics at the University of Munich, Germany, E-mail: felix.lauber@campus.lmu.de
 - This research paper was written for the Media Informatics Advanced Seminar on Interactive Surfaces, 2009

meaningful input-gestures for touch-sensitive displays. Consequentially a gesture - forming a semi-rectangle with ones hands to prevent the others from looking whats inside - was implemented. Also Sugimoto et al. base their Caretta-system on the fact, that collaborative working situations profit from the support of individual *and* group activities[14]. And another observation was important for their system: The problem in many collaborative working situations is, that a system having multiple simultaneous user-input but only one output-device (such as one display), users easily are getting confused because it is not easy to differentiate between the output of ones own action and those of other users. So in the Caretta system the attribute *shared spaces and private spaces* was implemented very consequently. In order to give a more detailed impression how the separation of shared spaces and private spaces can be implemented in an application, in the following chapters there are presented two example-projects: firstly the caretta-system which was just mentioned and secondly Coeno[3], which implements even three different spaces.

2.1 Caretta

The purpose of the caretta-system[14] is to support multiple users in an urban planning task: they can (re-)design a town by manipulating physical objects on a table and watch the effects of their changes in a computer-simulation. The system consists of three parts: small RFID-tagged objects like buildings which can be placed upon an RFID-reader equipped sensing board, an LCD-projector which augments the sensing board with the visualization of the computer-simulation and several PDAs representing each user's private space.

2.1.1 Shared space - the sense board

The sense-board realizes the shared space of the system. Standing around it, users can add, remove small RFID-tagged model-buildings or change their position. Each time the scene has been changed, the simulation system evaluates the effects of the changes on several parameters like the town's population, revenue and expenditure. Those effects are visualized and projected as an overlay onto the senseboard, so the users can immediately see and evaluate the effects of their changes.

2.1.2 Private space - the PDA

By using his PDA each user has the possibility to realize his own private space. The PDA offers two modes: the personal mode and the shared mode. Once a user has entered the shared mode, each change which has been made to the system by other users appears synchronously on his PDA-screen. The personal mode realizes the pure private space. Users can change their personal copy of the current system-state without affecting the global system-state. After having manipulated the objects on the PDA, the user can see which effects the change would have to the system if it was applied. So the user can evaluate his plans before presenting them to the others.

2.1.3 Results

The user studies for Caretta showed that all users used the shared and private spaces as they were meant. The PDAs were used to devise and evaluate their own plans, while the sensing board was used for discussion and negotiation with other users. Another interesting observation during the user-studies of the caretta-system was, that the PDAs were not only used for private work in the strict sense of one person using the system but also for forming subgroups of two and using it for temporary subgroup-collaboration (in the sense of a semi-private space). That means that private spaces need not necessarily refer to a single user-mode, but also to a subgroup of users sharing the same kind of ideas and objectives.

2.2 Coeno

Another example for an application, where the separation of differently accessible spaces plays an important role, is Coeno[3]. Coeno is an application for storyboard design, which is a challenging task demanding a high level of collaboration. In Coeno there exist three

different spaces: the private space, the design space and the presentation space.

The presentation space is a display projected to a wall in the meeting room, where the work can be presented to all people in the room. Each member of the working group brings his own laptop, representing his private space. The private space is the only place, where a team member can create new content (for example images, sequences or scribbles) or modify it. The laptops are connected wirelessly to a display server, so once having created new content the user can drag it with his mouse to the design space. The design space is only used for structuring information and discussing it within the group. When a user wants to edit content, he must drag it back to his private laptop, modify it, and copy it back to the design space. On the design space people can move the content, scale and rotate it on the table, all users at the same time. When they decide to bring certain content into a particular order in the storyboard design they can move it into a fixed position in the storyboard-timeline of the presentation space (*see figure 1*). In the user studies for Coeno, a lot of users liked the idea of



Fig. 1. Different spaces of Coeno.[3]

having different work spaces for different tasks. An interesting finding was, that half of the users wanted a direct communication link to the other users for sharing content in a direct way (see also chapter 4).

3 FAST, DIRECT AND CONCURRENT INTERACTION BY ALL CONTRIBUTORS AT ANY TIME

When performing a collaborative task, it is necessary for avoiding production block, that every user has the possibility both to see the presentation of whats happening on the workspace and to join in the task whenever he has the need[4]. Basically there are two possibilities to fulfill this requirement of *fast, direct and concurrent interaction by all contributors at any time*:

- Each user interacts with his own device. Either this device is only used as input unit and the output of the system is presented in a way all of the users are able to access; this solution is implemented for example in Mobile Notes[1]. This system can be used for example in an educational environment like a classroom. By using PDAs as input devices, pupils have the possibility to contribute texts, scetches or votes to a central server, in order to display and arrange these products on an electronic whiteboard. The alternative is, that this device is both input and output unit and the users are connected with each other using a network infrastructure. This form of remote collaboration is realized for example in the CommChair[13] and the ReactTable[7]. Note that because of the essential necessity of a network connection between all devices this solution is closely related to chapter 4.
- One device, used as input and output unit, is shared by all users. This solution may be realized by using a multi touch-table (for example in Hilliges' digital brainstorming[4]) or a tangible user interface (for example in Caretta[14] or Senseboard[6]). The Lumisight Table[9] tries to solve one problem of horizontally

aligned displays: only a part of the users can see the display in the right orientation.

In the following chapters there are presented two example-projects, one for each of the solutions just introduced. The reacTable[7], which is presented at first, actually implements both of the possibilities of concurrent interaction; because of its characteristic remote collaboration mode the explanation though concentrates on this aspect of the project. The Lumisight Table, which is presented afterwards, is representative for the second possibility of concurrent interaction. Although the idea of using one horizontal display which is shared by all users is quite obvious, the solution of the orientation problem is very interesting and helps to fulfill the requirement of concurrent interaction.

3.1 reacTable

The general idea of the reacTable is to provide an electronic music instrument based on a tangible interface. It consists of a translucent round table, small objects representing components of a synthesizer which can be placed on the surface of the table and a video camera which detects the constellation of the objects on the table and their orientation. Users can grab the objects, move them and produce different sounds by arranging different object-constellations. The reacTable is implementing both methods of collaboration mentioned above: local collaboration where users are standing around the multi touch table but also a remote collaboration mode to which we want to go into detail. In the remote collaboration mode users have the possibility to connect two or more tables placed at distant locations. Once the tables have been connected, users have the impression of sitting together at one single table. Changes which have been made by one user appear as virtual object projection on the display of the other users. Both the real physical objects on the surface of the table and the virtual projected ones behave the same way, even though the users cannot manipulate the virtual objects. In order to explain the remote collaboration mode of the system, Kaltenbrunner et al. introduced an example scenario where two players, one of them being in Barcelona, the other one in Linz, are working together. According to this scenario, the player in Barcelona for example can place a sound generator object on his table. The system's sensor component recognizes the object, locates it on the surface and projects its graphical representation underneath the physically present object. By transmitting the relevant data, the same graphical representation of the object is simultaneously projected at exactly the same position on the table in Linz, where it has also been located at the table in Barcelona. Now the player in Linz places for example a sound effect on the table and after the same process of recognition, transmission and display, the sound effect appears projected on the table in Barcelona[7]. By just transferring control data like the object's id and its orientation but no audio data at all, the remote collaboration has a low latency. The audio data is locally produced at each user's table by interpreting the received control data.

3.2 Lumisight Table

Using a multi touch table is a very popular decision, as they realize the possibility of multi-user interaction in a way, which is very natural and comfortable. The problem with multi touch tables is that, even if they realize the input-direction very well, only a part of the users has the correctly oriented view and for others it appears in the worst case up side down. Matsushita et al. enunciate two conditions for a satisfying collaborative environment: first, the *sharing of nonverbal modalities* like gestures or eye-contact among the users and second, the *equal access* by all users to the information shown on the display[9]. Therefore Lumisight was developed, a system which is able to support four users equally with the same view on the table. By using an innovative display technology this is possible without forcing the user to wear any devices disturbing the communication.

3.2.1 Hardware

The second condition, the equal access to the information displayed by all users can be satisfied by using a special hardware setup (see figure 2). It is the combination of the four projectors, the mirrors and

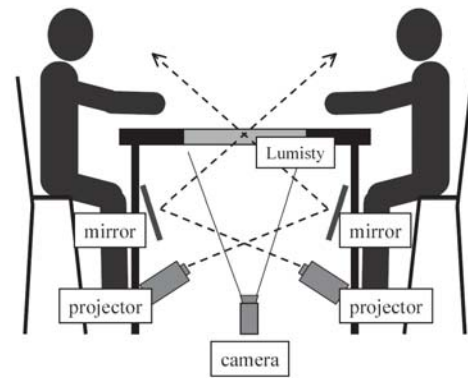


Fig. 2. Technical overview of Lumisight.[9]

the two orthogonally aligned layers of Lumisight film on the top of the table, which provides the functionality of projecting differently orientated images to each user. The Lumisight film is able to filter out an image according to the user's direction (see figure 3) so using two orthogonally aligned layers of it makes it possible to support the four user setup.

For the purpose of capturing the user's input for example with hand gestures or objects the cameras are installed inside the table. As the Lumisight film is translucent to light coming from the vertical direction, it is possible for the cameras to capture objects upon the table.

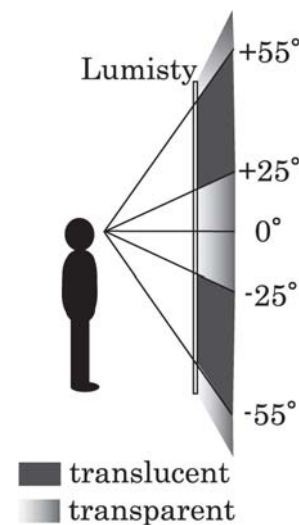


Fig. 3. Optical property of Lumisight film.[9]

3.2.2 Software

Although the first condition, the possibility of sharing nonverbal modalities, is not violated by the system (disturbing devices as HMDs are not necessary), it is also not satisfied by the table so far. If every user works with a differently oriented copy of the same image, sometimes it is not an adequate basis for communication as for example a spot being pointed at by some user is not the same spot on the view of the other users. Therefore a complex catalog was developed, separating the information being displayed into different categories. According to this catalog, some information need to be rotated and re-oriented (as for example annotations to important spots displayed in a sub window) while other must have the same location, rotation or even both for all users.

It is also remarkable that by implementing a method for displaying different types of content differently to the users, another approach

for differentiating between shared and private spaces (see chapter 2) is found.

3.2.3 Results

A preliminary studies of the LumisightTable showed that the table indeed can improve the user's comprehension of the system. The setup for the study was as follows: an instructor and two participants use the system while all of their utterances are recorded. The same experiment was made using a table without the multi-user support of Lumisight and the number and quality of utterances was compared. The observation being made, was that using the Lumisight table decreases the utterances of the instructor concerning coordination and problem solving while the utterances of the participants concerning topic-related discussion increase.

4 DIRECT COMMUNICATION LINK OF PARTICIPANTS AMONG EACH OTHER

In Caretta (see chapter 2.1) users had the requirement of forming subgroups and developing ideas within a semi-private space. In this case, users of one subgroup shared one device and one display by surrounding one PDA. But in some application scenarios it can be necessary for the users not only to communicate - verbally or non-verbally - but also to exchange digital information. This can be necessary, for example, when the whole group splits off, every user works on his own task and after some time the group rejoins and they want to merge their results[13]. Another application, where half of the test persons in the user study had the requirement of a direct communication link to the other users is Coeno(see chapter 2.2). In this case it was too cumbersome first dragging content to the design space so that afterwards another user could drag it from there to his own private space.

Also the requirement of supporting a concurrent multi user input within the subgroup can easily be supported by establishing a communication link between user's devices. In the iLAND environment[13], for example, there is the possibility of using the CommChairs in a shared display mode so several users can work on one virtual display while they do not really have to share it physically. The last aspect could also be advantageous to allow very large sized collaborating groups where there would not be enough space for every member surrounding one single device.

Another case, when digital data transformation between the users is a useful property is the AR-based meeting environment MagicMeeting[5]. Using PDAs, this system supports easy to use communication links both within the whole group and between the presentation area and the individual devices each user brings to the meeting. These examples reveal the need of the integration of the next attribute, a *direct communication link of participants among each other*. In the following chapters we will have a closer look at two projects. The first one is the iLAND-project, which introduces not only an application but a whole environment facilitating collaboration and creative tasks and which is based on the idea of every element being connected with each other. The second project which is presented in detail is MagicMeeting, as this application also necessarily needs network support to make the model exchange possible.

4.1 iLAND

When iLAND was developed, the overall idea was to design an all-embracing environment being able to support a couple of common collaboration scenarios. For that reason several pieces of furniture were developed, all designed to support collaborative activities in those scenarios with the help of integrated digital devices. Additionally a software infrastructure was developed, which allows communication among the roomware components within a network infrastructure. Only by using this possibility of communication the full benefit of each of the components and the whole environment is available.

In the next sections there is given a short overview of each of the roomware components in order to explain their functionality and how they can be used together.

4.1.1 DynaWall

The DynaWall is a huge touch-sensitive display, which covers a whole side for example of a meeting room. As common interaction techniques, like dragging an object from one side of the wall to the other, are not easily realized because of the dimension of the wall, new mechanisms had to be developed. The 'drag and put' for example provides a way to put an object at one side of the wall, go to the other side without having contact to the wall, and put it somewhere on the other side. The 'shuffle' feature allows the user to throw an object across the wall which can be caught by another team member on the other side. Additionally it is equipped with a network interface so it can be used together with the other roomware components.

4.1.2 CommChair

The CommChair is a chair, which either is equipped with a computer integrated in a swing-up desk or provides a docking station where users can connect their own laptop. Like the other roomware components it is equipped with a network interface and supports connectivity to other devices in the room. When users want to form a subgroup, for example, they can push together a couple of CommChairs, the chairs detect the proximity to each other and enter a shared display mode. In this mode users have the impression of working together at a logically homogeneous workspace. For realizing this concept (which works with all of the roomware components) a special software infrastructure was developed, called BEACH.

4.1.3 InteracTable and the Passage-Concept

The InteracTable is a interactive table, which provides a horizontal touch sensitive display which can be used by two to six users standing around it. Shuffling and rotating objects is possible by simply using the touch-sensitive display and a wireless keyboard provides a possibility for text input.

The Passage-concept is an interesting concept which can be used with the table. The concept implements a way that each detectable object, like cups, pens or rings for example can be associated with a piece of information and when the object is placed on the table, the information is displayed. Note that the object does not really carry the information, but only provides a link to it. The implementation in iLAND uses the weight of the item for object recognition.

4.1.4 Using them all together

Imagine a collaboration scenario in a big meeting room well equipped with the roomware components introduced in this chapter. A collaboration team has the possibility to split up into subgroups, each subgroup working on another aspect of the task. One subgroup for example structures information using the InteracTable while another subgroup is using the CommChairs with the shared display mode. With their private Laptops and the connection between them, they can either produce new aspects of the task or load additional information from the laptop into the environment to provide the structuring team at the table with new data. The results of their subgroup work can be brought to the DynaWall where they can be presented to all people in the room or - after reunion of the subgroups - the whole team can discuss and merge them.

4.2 MagicMeeting

The idea of MagicMeeting is to improve a meeting situation where several users bring 3D models of objects they want to present and discuss with the others. Therefore a meeting situation was constructed where all users wear head mounted displays(HMDs). In the middle of a table, which is surrounded by the users, there is a turnable, flat device, called 'cake platter', which is used as the central presentation space for 3D objects. As the main interaction device, each user has a PDA containing the markers of the models he wants to present. Users can browse in their personal selection of markers, increase their size to full screen and see the augmented model represented by the marker. When he wants to present a model to the other users, he simply moves his PDA next to the cake platter and the selected model in his PDA is transferred to the cake platter while the model formerly being on

the cake platter moves to the PDA. After this exchange of models, the marker is displayed on the cake platter and so can be examined by the other users. As the cake platter is not fixed on the table, users can change its position, turn it or tilt it for better model examination. When they want to see the lateral cut of the model and see its interior, they can hold a marker-equipped clipping-plane inside the model. When a user wants to have a model of another user on his own PDA, he asks for it and the model can be sent to him by the other user via the infrared interface of their PDAs.

The informal evaluation of the system showed that people liked the interaction with the augmented 3D models. Especially that every user has his individual perspective of the model and the flexibility of the cake platter were evaluated positively. Users thought that the interaction devices like the clipping plane and the PDAs could be used intuitively. Problems of the system were originated technologically: reflections of light on the display of the PDA or the clipping plane interfered the marker recognition of the system.

An interesting examination would be the effect of the HMDs on the user's communication and group interaction. McAtamney et al. found out that wearing HMDs can be disturbing in informal conversations[10].

5 PHYSICALITY OF OBJECTS

Unlike the three chapters before, the following two chapters do not explicitly refer to collaborative working scenarios but can enhance creative processes. For this purpose it is important, that the system interaction is intuitive and natural. A very important attribute in this context is the *physicality of objects*. Hilliges et al. propose in their design considerations physicality as being relevant especially for applications, which are similar to their touch-sensitive display based digital brainstorming system[4]. They introduce two kinds of physicality: *pseudo-physicality*, which means the use of digital metaphors for existing objects, and *meta-physicality*, which means object-behaviour or -appearance being unrealistic and only possible in the digital world but predictable and logical though. A project that can be mentioned in the context of meta-physicality for example is a novel display metaphor called *Display Bubbles*[2]. By using no rectangular objects but shapes which can be formed freely, the existing space of the display can be used more effectively. This is for example the case, if working with a tabletop setting and objects on the table are constraining the extension-space for displayed objects.

Other Projects, like for example Senseboard[6] and Caretta[14] blend real physical objects with digital interaction. When users handle with real objects or at least metaphors being derived from real objects, they find it easy to learn the appliance of the user interface[4].

In the following chapters, two projects are presented in a more detailed way. The first one is the brainstorming application of Hilliges et al. where the two aspects of physicality are also explained more thoroughly. The second one is the Senseboard, which exemplifies the usage of real physical objects in a digital application.

5.1 Digital brainstorming in an instrumented environment

As mentioned above, Hilliges et al. implemented certain aspects of physicality in their digital brainstorming system[4]. They differentiated between *pseudo-physicality* and *meta-physicality*, both shaping out as being useful for his application. In this chapter we like to give a short overview of the application itself and a short description of both aspects of physicality.

5.1.1 Brainstorming application

The main idea of Hilliges' digital brainstorming system was to carry the traditional idea-card-based brainstorming technique to an instrumented environment. This environment consists of two main parts. One part is a multi touch-table being the area where users can create their idea-cards. Once having created an idea-card, users can move them upon the table, edit them or change their size. The other part of the working-environment is a interactive display wall, placed next to the table. On this wall, each idea-card appears immediately after creation, is oriented upright and can be moved or structured by forming

groups. These two spatially separated system-components represent two important phases of the brainstorming process: the table is used for the generative phase, while the display wall is used for the structural phase.

5.1.2 Pseudo-physicality

The first aspect of physicality is the *pseudo-physicality*. This means, that objects existing in traditional brainstorming are used as a metaphor for the equivalent ones in the digital version. So users deal with digital idea-cards, which are very similar to post-it-notes they are used to. Users can write on them with a pen and move them around. In conjunction with the touch-sensitive surface of the table users have the feeling that they can interact with the system just like they know it from real world.

5.1.3 Meta-physicality

The second aspect of physicality is the *meta-physicality*. Hilliges et al. think, that implementing metaphors from the real world can sometimes be insufficient and specific affordances of digital media can also be helpful for supporting the users interactions. He made the experience, that even if this behavior is somehow unrealistic users accept it, it just has to be predictable and explainable. In the brainstorming-system there were implemented two examples of such behavior: the automatic re-orientation of the idea-cards on the wall and the possibility of skidding them around on the table. The technique of skidding means, that users can give an impulse to a card by doing a quick movement with the pen and cause that the card is sliding to another user. This implicitly gives the users the idea of passing the card over to another user which can support social communication.

5.1.4 Evaluation

The most important result of the evaluation of the digital brainstorming application was, that is is not disruptive. So even if the communication of the candidates decreased slightly in comparison to the paper based brainstorming, the number of ideas was the same and their quality even increased.

So even if there are no advantages of the digital version concerning the direct collaborative work session, electronic brainstorming still can profit from the advantages of using digital systems in general. Exporting and saving the results of work, the possibility of entering earlier system states and the combination with automatic information browsing- and structuring-tools (see chapter 6) could generate the real advantages of such a system.

5.2 Senseboard

Concerning the field of application the Senseboard-project[6] is quite similar to the brainstorming application just presented above. But instead of using a large display and gesture input, the idea of Senseboard concerning the technical realization is to use a tangible user interface. The Senseboard-system consists of a large vertical panel equipped with RFID readers. The surface of the panel is marked with a rectangular grid and users can place small rectangular 'pucks' into the cells of the grid. Containing a passive RFID tag, the pucks can be recognized by the board including their identity and location, once they are placed on the board. The system supports two other possibilities of command input: the user can send a command either by pressing the surface of the puck or by briefly placing a special 'command puck' over the puck. The system UI is projected onto the senseboard using a video projector.

5.2.1 System interaction

As the system is using physical objects as input devices, interaction is very natural: the user sees the item he wants to interact with, he can grab it or move it. As mentioned before, there exist two categories of pucks. The data puck is the puck carrying the pieces of information the user wants to structure and organize. The command pucks, in contrast, are different in shape and size each one representing a special command. By placing for example the 'view details' puck over a data puck, some additional information of the data puck is projected onto

the 'view details' command puck. Another interesting command is the 'group' command: by placing a special 'group' puck over the topmost of a couple of stacked data pucks, the user logically puts them into one group (see figure 4). Other command types like 'copy', 'link' or 'type-in' (which provides a possibility to add typed text to a data puck) are also provided.

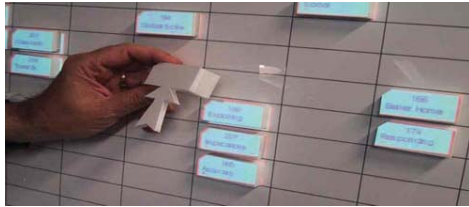


Fig. 4. Using the group command puck.[6]

5.2.2 Evaluation

The assumption is, that the system has two main advantages to the non-digital sticky paper alternative:

- By doing a structuring task, for example organizing a schedule, there is the possibility of items conflicting with each other. In the digital version, the system detects any conflicts and visualizes them on the board.
- After having accomplished the task, the user has the possibility to export the final arrangement of items so that it can be processed by another software. Using sticky papers, the user would have to transcribe the results of his work.

When executing the user study especially two comparisons are interesting: the one with the sticky paper notes version of the system and another one, using a more conventional GUI with a whiteboard and a none marking pen instead of the pucks. The result of the user study of Senseboard was that the Senseboard seems to perform slightly better than alternatives using sticky paper or the GUI. Users generally liked the idea of manipulating real physical items. The disadvantage of the system design was, that by interacting with the system, sometimes users covered the projection with their body moving in front of the board. But the attribute of physicality seemed indeed to bring advantages, especially in comparison to the GUI version.

6 AUTOMATED SUPPORT OF BROWSING AND STRUCTURING INFORMATION

As mentioned in the section before, also this attribute does not explicitly enhance collaborative working processes, it is rather extremely important for supporting creativity. Creative work very often includes the subtask of collecting and structuring information, as for example in a phase of brainstorming[4]. Collecting information usually happens in a very early phase of creative work while structuring it normally occurs in a later phase[12]. Nevertheless the two tasks are somehow connected to each other, as collecting hundreds of pieces of information does not make sense without structuring them and the other way round. However both of the tasks can be supported automatically and thus help the user in fulfilling them, which leads us to the next attribute *Automated support of browsing and structuring information*.

The first project presented in this chapter is called *Implicit brushing and target snapping* and supports information structuring tasks[15]. The second project presented in this chapter is called *CombinFormation* and supports the user in collecting relevant information[8]. Another project that can be mentioned in the context of finding the right information, is the software framework implementing the features IAM-eMMA and EVIDII[11].

6.1 Implicit brushing and target snapping

The idea of the work of Sun et al. was to implement a system which assists the user in grouping pieces of information without being intrusive[15]. This can be useful for example in brainstorming applications, when dozens of little pieces of information need to be structured. Therefore the implicit brushing technique was developed. This feature is triggered by the user when he creates a group of items on the screen by placing the data objects in a way that they are overlaying each other. The system automatically analyzes group member text and metadata and proposes new group member candidates. This happens by giving their frame the same color as the group member's and by adding an arrow to each candidate pointing in the direction of the group. The implicit brushing technique thereby does not only take the data objects on the screen into account but also data objects being located in a background database. For not confusing the user the different relevance levels of the data objects is mapped to their size, so that objects retrieved from a background database are smaller than objects the user has created. When the user wants to add a candidate to the group, he clicks at the arrow of the candidate and the item automatically moves to the group. This feature is called target snapping and was implemented as drag and drop operations usually do not work well on large displays. For an overview of how brushing and target snapping are used together see figure 5.

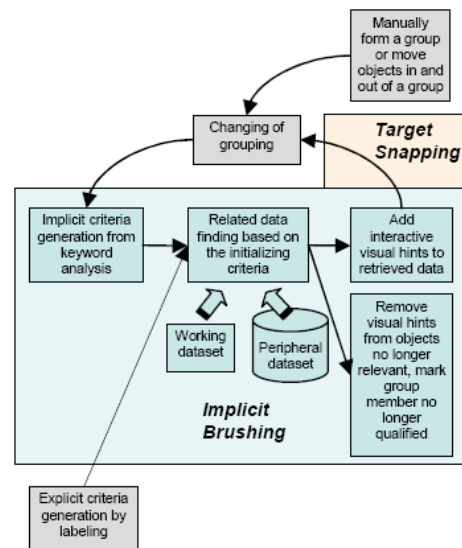


Fig. 5. Brushing and target snapping.[15]

In an informal user study the new techniques were tested in two application scenarios. In the first one, an electronic brainstorming application, both the implicit brushing and the target snapping helped the users very well in grouping the data objects. Also in the second scenario, structuring a collection of photos manually tagged with metadata, the application performed very well and supported the users grouping actions.

6.2 CombinFormation

Browsing for information is the main benefit CombinFormation[8] provides to the user. The main idea of the system is, that the cooperation of the system and the users interaction increase the effectiveness of information browsing in the web. In order to explain the collaboration between the user and the system, in the following both the system's initiative and the user's initiative are described in detail.

6.2.1 User's initiative

The system provides three different modes that can be selected for initial system launch. The user can either open an existing composition

for further collecting and composing (*re:open*) or he can give a number of seeds (initial browsing keywords) to the agent (*re:mix*) or, in the third mode, he can work with a predefined set of seeds provided by the system (*re:collection*).

Afterwards the user can work with surrogates, which are image or text objects linked with resources on the web. These surrogates can be created by the user by dragging objects from his web browser to the application space. These objects can be images, links, files or simply text. By using a toolbar, the user can give more or less relevance to single surrogates, edit them manually, remove them or manipulate their appearance.

6.2.2 System's initiative

The system tries to support the user's browsing by constantly presenting him new candidates related to his search. While the user interacts with the system, the application permanently evaluates the user's input and modifies its search. An overview how the user's initiative and the system's initiative are working together is given in figure 6.

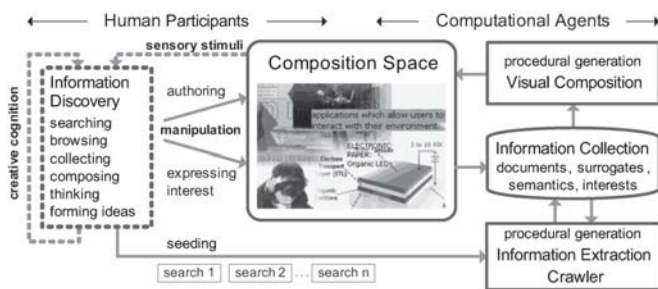


Fig. 6. Collaboration of the user and the system.[8]

6.2.3 Evaluation

In the evaluation studies of the system users thought that the system provides help in developing new ideas. They could be very creative in combining different perspectives of the information provided by the system. They also liked the system itself with its image and text representations of web content. Searching for very specific content in contrast (for example one particular image) would be easier with Google. CombiFormation was initially designed only to support a single user mode. But when users were asked to propose new features for the system they also suggested to integrate collaboration modes. A collaborative composition space, for example running on a multi touch table, could be an interesting feature for future development of the system.

7 CONCLUSION

In this paper we showed the effects of the integration of technology in creative working processes. On the one hand there are no significant disadvantages for interpersonal communication activities, as for example Hilliges et al. have ascertained[4]. On the other hand some sub-tasks of creative working can be simplified and successfully supported as for example in CombiFormation[8] or in 'Implicit Brushing and Target Snapping'[15]. Also distant collaboration as in reacTable[7] and the additional information overlay on tangible interfaces as the Senseboard[6] and Caretta[14] are in this development only possible with the support of computer-based technology.

In this paper we presented furthermore five design suggestions for applications, which support creative working processes by making use of interactive displays. The first one was *separation of shared spaces and private spaces* which is necessary for supporting the different working phases in a collaborative, creative task. The second one, *fast, direct and concurrent interaction by all contributors at any time* has to be supported by the application to avoid a production block of any user due to the fact that he wants to contribute something but cannot. Establishing a *direct communication link of participants among each other* is the third important feature we described in the paper; we gave several examples of projects and tasks where this can be useful or even

indispensable. The fourth attribute, *physicality of objects*, leads generally to a better understanding and acceptance of the application and makes interaction more natural; Jacob et al. for example showed, that this improvement leads to better results in creative tasks. The last attribute, *automated support of browsing and structuring information*, can improve and fasten problem solving and creative tasks in supporting one specific subtask of many creative tasks: the collecting and structuring of related information.

Considering Shneiderman and the four phases of creative processes, they can be assigned to them as follows. For the first phase, the collection of information, the *automated support of browsing and structuring information* is extremely useful as one can see for example in the CombiFormation Project (see chapter 6.2). The *direct communication link of participants among each other* can support the second phase, where the user wants to consult peers and mentors. Especially when working in a group which is not in the same room or when there is the necessity of exchanging digital information, this attribute would be indispensable. The attribute *separation of shared spaces and private spaces* can be assigned to the third phase and especially to the action of exploring different solutions. Particularly in collaborative processes, when several users are having several ideas the same time, it is extremely helpful to give them room to try out and evaluate their ideas on their own before presenting them to the group. Also the attribute *fast, direct and concurrent interaction by all contributors at any time* is very useful for the phase of creation if it was interpreted in a way that would be suitable for collaborative processes. The attribute *physicality of objects* would also fit into the phase of creation as it supports generally the interaction and the comprehension of the users. It is an interesting fact, that none of the applications presented in this paper implemented the fourth aspect, the dissemination of the results, in a remarkable way. Probably the reason for this is, that it is not very obvious to support this last step of creativity by using interactive displays.

REFERENCES

- [1] L. Bollen, G. Juarez, M. Westermann, and H. U. Hoppe. Pdas as input devices in brainstorming and creative discussions. *Wireless and Mobile Technologies in Education, IEEE International Workshop on*, 0:137–141, 2006.
- [2] D. Cotting and M. Gross. Interactive environment-aware display bubbles. In *UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology*, pages 245–254, New York, NY, USA, 2006. ACM.
- [3] M. Haller, M. Billinghamurst, J. Leithinger, D. Leitner, and T. Seifried. Coeno: enhancing face-to-face collaboration. In *ICAT '05: Proceedings of the 2005 international conference on Augmented tele-existence*, pages 40–47, New York, NY, USA, 2005. ACM.
- [4] O. Hilliges, L. Terrenghi, S. Boring, D. Kim, H. Richter, and A. Butz. Designing for collaborative creative problem solving. In *C&C '07: Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition*, pages 137–146, New York, NY, USA, 2007. ACM.
- [5] M. W. H.T. Regenbrecht and G. Baratoff. Magicmeeting: A collaborative tangible augmented reality system. *Virtual Reality*, 6(3):151–166, 2002.
- [6] R. J. K. Jacob, H. Ishii, G. Pangaro, and J. Patten. A tangible interface for organizing information using a grid. In *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 339–346, New York, NY, USA, 2002. ACM.
- [7] M. Kaltenbrunner, S. Jorda, G. Geiger, and M. Alonso. The reactable*: A collaborative musical instrument. *Enabling Technologies, IEEE International Workshops on*, 0:406–411, 2006.
- [8] A. Kerne, E. Koh, S. M. Smith, A. Webb, and B. Dworaczyk. combiFormation: Mixed-initiative composition of image and text surrogates promotes information discovery. *ACM Trans. Inf. Syst.*, 27(1):1–45, 2008.
- [9] M. Matsushita, M. Iida, T. Ohguro, Y. Shirai, Y. Kakehi, and T. Naemura. Lumisight table: a face-to-face collaboration support system that optimizes direction of projected information to each stakeholder. In *CSCW '04: Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 274–283, New York, NY, USA, 2004. ACM.
- [10] G. McAtamney and C. Parker. An examination of the effects of a wearable display on informal face-to-face communication. In *CHI '06: Pro-*

- ceedings of the SIGCHI conference on Human Factors in computing systems*, pages 45–54, New York, NY, USA, 2006. ACM.
- [11] K. Nakakoji, Y. Yamamoto, and M. Ohira. A framework that supports collective creativity in design using visual images. In *C&C '99: Proceedings of the 3rd conference on Creativity & cognition*, pages 166–173, New York, NY, USA, 1999. ACM.
 - [12] B. Shneiderman. Creating creativity: user interfaces for supporting innovation. *ACM Trans. Comput.-Hum. Interact.*, 7(1):114–138, 2000.
 - [13] N. A. Streitz, J. Geißler, T. Holmer, S. Konomi, C. Müller-Tomfelde, W. Reischl, P. Rexroth, P. Seitz, and R. Steinmetz. i-land: an interactive landscape for creativity and innovation. In *CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 120–127, New York, NY, USA, 1999. ACM.
 - [14] M. Sugimoto, K. Hosoi, and H. Hashizume. Caretta: a system for supporting face-to-face collaboration by integrating personal and shared spaces. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 41–48, New York, NY, USA, 2004. ACM.
 - [15] X. Sun, P. Chiu, J. Huang, M. Back, and W. Polak. Implicit brushing and target snapping: data exploration and sense-making on large displays. In *AVI '06: Proceedings of the working conference on Advanced visual interfaces*, pages 258–261, New York, NY, USA, 2006. ACM.
 - [16] M. Wu and R. Balakrishnan. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *UIST '03: Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 193–202, New York, NY, USA, 2003. ACM.

Applications for Interactive Public Displays

Sebastian Löhmann and Bettina Conradi

Abstract—More and more public displays can be found in our daily life. But most of them are still not interactive. The users can only look at and remember the contents, but they have no possibility to capture them to take a closer look later on. However, there are many research projects with the goal of making public displays interactive. The idea of this paper is to provide an overview and a categorization of relevant ideas and approaches. We developed five categories which are supposed to describe different ways in which users can interact with public displays and thus handle the shown contents: View, Download, Influence, Add and Change. With the example of different research projects we show benefits and problems of each category and discuss possible solutions. Future researchers can allocate their planned applications to one or more of the interaction categories and thus find facts they need to be aware of while dealing with the topic of interactive public displays.

Index Terms—interactive public display, public displays, interaction categories, user needs

1 INTRODUCTION

Nowadays public displays can be found in many places. Whether squares, shop windows or subway stations, public displays become more and more popular. People use them to gather information about the weather forecast or latest news, companies advertise their products and passengers fill their waiting time with cartoons and music videos. But are those displays really interactive? Let's have a look at the Oxford English Dictionary: "[interactive:] involving direct communication both ways, between the computer and the person using it" [1]. Obviously, the communication from a person to the public display is rather limited so far. The current form of interactivity is very weak.

But why would that be important, anyway? To communicate and interact with something is for most people more interesting than simply watching it and being passive. And from the public display's owner's point of view: people will more likely be interested in the presented contents, if it increases their creativity and stimulates communication and thus makes them active.

Technologies, which developed during the last years, allow different kinds of interaction between public displays and their users. While for example touch screens are a good way to attract people, they are not sufficient for many scenarios. Public displays are oftentimes not reachable and if they are, they might be protected because of possible vandalism. On the other hand, technologies like mobile phones, with features such as SMS, camera or WLAN, provide perfect conditions for making public displays interactive.

Interactivity makes it possible for viewers to actually use the shown contents instead of only looking at them. This fact raises another question: What can people actually *do* with the displayed material? To answer this question, we created five categories, which are supposed to describe different ways of dealing with those contents. In this paper we want to discuss current research projects, which aim to make public displays interactive, and classify them with the help of our categories. Relevant ideas will be summarized, we will take a look at related user studies and try to discuss the projects' potentials.

2 INTERACTION CATEGORIES

2.1 View

Talking about interaction between public displays and their users, the most simple action for people is to view the provided contents. It is the weakest form of interaction, because the viewer is rather passive. He

can only look at the display, but he can neither obtain the contents (and thus tend to forget them) nor influence what the screen will display next. For most people, viewing content is nothing new, as they are used to watching television.

2.1.1 Public Displays in Subway Stations

A current example are the big screens in many subway stations. They show the news, weather forecasts and little cartoons. Additionally, more and more displays showing timetables are installed. Obviously, the purpose of this service is to entertain and inform the people waiting. From our own experience on a usual work day we can define a waiting time between 2 and 10 minutes, depending on the day time, before we can enter a train. Isn't it enough time to perform some interactions with the public displays? Isn't it enough time to take out our mobile phone to send a SMS or take a picture? But besides viewing the (by all means useful) information, there is no interaction possible. The displays in subway stations will serve as a running example of this paper. It will be relevant for all categories and can show the way towards *interactive* public displays in an illustrative manner.

2.1.2 GroupCast

A research project fitting into the category "View" is GROUPCAST [8]. This application aims to initiate casual interactions between the people nearby. It uses infrared badges and sensors to recognize who is near a public display. Then it tries to find out common interests of those persons and displays related information. In that way the system is supposed to help people who don't know each other very well to start a conversation about the viewed topics. Again there is no direct communication between the public display and the user, but a positive aspect is the try to create interactions between both of them. Unfortunately, GROUPCAST didn't work out. The problem was the gathering of information about the users. To find two persons' common interests (which have to be good enough to encourage a conversation) turned out to be very difficult. The researchers tried to develop web-based profile forms, but they were so large that users could not afford to take enough time to fill them out. A concern of the people anticipating in the study might have been the privacy aspect (unfortunately, the paper does not provide information on this topic). Who is willing to give out his personal information details while he knows that they will be published on a public display later? We will see later on, that for most people this has a discouraging effect on the use of public displays. Researchers should be aware of this problem while they are dealing with sensible data.

Both of these examples have two things in common: They aim to show useful information for interested people, but they do not allow any direct interactions by the users. As mentioned before, people forget relatively fast, what they have just seen. How can that be prevented?

- Sebastian Löhmann is studying Media Informatics at the University of Munich, Germany, E-mail: sebastian.loehmann@campus.lmu.de
- Bettina Conradi is a researcher of the Media Informatics Group at the University of Munich, Germany, E-mail: bettina.conradi@ifi.lmu.de
- This research paper was written for the Media Informatics Advanced Seminar on Interactive Surfaces, 2009

The logical next step would be the ability to retrieve the information found on a public display. How this could be achieved will be shown by the next category, "Download".

2.2 Download

Mobile devices have become more and more popular. In Germany, more than 90% of the people own a mobile phone. They have become an important part of the daily life, not only for the purpose of communication. Mobile phones are much more than that: built-in technologies like cameras or Wi-Fi make picture/video capturing or internet usage possible. Office applications, integrated MP3 players and multimedia browsers are just a few more features of modern mobile phones. Using these opportunities, ideas came up of how to capture contents from public displays.

A relevant research project is Shoot & Copy by Boring et al. [2]. This paper describes the first active form of interaction between public displays and their users we want to present: Downloading provided information. The most important feature of the mobile phone used by this application is the integrated camera. Using the 3.2 mega pixels by Sony Ericsson's K800i, taken pictures have a quality good enough to process them and retrieve contained information. The basic idea is the following: a large screen shows a number of icons, each representing a media file such as audio, video or documents. The user can select a desired icon by taking a picture of the relevant part of the screen (see figure 1). In the next step, the picture is sent to the computer controlling the public display. Due to a relatively big distance between camera and display (they assume a minimum of 50 centimeters), in most of the cases the photograph will contain more than one icon. This problem is solved by an image processing software, which is able to identify the correct icon by comparing the picture to the actual display content. The software chooses the icon closest to the picture's center. Now, the user has two options: he can either ask for a URL corresponding to the media file, or he can download the actual file directly to his phone. Using either way, it is now possible to easily "remember" interesting contents. The first step of interacting with a public display is made! If the user finds some time to look at the retrieved information in detail, he can for example connect his mobile phone to a personal computer. To make this possible, Boring et al. used the Bluetooth-ability of the K800i. An application running on the computer can now receive the list of URLs gathered by the user. It is now possible to show or play back the media files in a common browser.

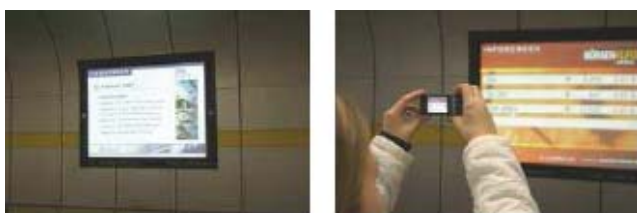


Fig. 1. Shoot & Copy: Retrieving information from displays in a subway station [2]

An user study involving 28 participants has been part of the project. The main purpose of the evaluation was to find out if the application was easy to use (even by people who are not familiar with mobile phone cameras) and if the users could imagine to actually use the application in certain scenarios. The results were throughout positive: 82% of the participants stated that they would use the system if it was available. Interesting to them was especially our scenario of public displays in subway stations mentioned before. The time before the arrival of the next train would obviously be enough to download some files of interest. This is the first step on our way to *interactive* public displays. The study also attested the application one of the most important factors: it was easy to use. All participants were able to select the correct icon with their first try.

But there were also two negative side-effects. In two cases, the application did not work right away because of instable Bluetooth connections. However, Boring et al. stated that mobile phones are increasingly equipped with wireless LAN modules, which promise a rather reliable connection compared to Bluetooth. The Wi-Fi ability can also solve the second problem: Although no participant of the user study felt that the time between taking a picture and receiving was too long, an average time of 9.2 seconds sounds still long. But wireless LAN connections are not only more stable than Bluetooth connections, but also faster. Hence Shoot & Copy can only be improved.

2.3 Influence

So far, being people in a subway station waiting for a train, we can do two things: we can pass the time reading information provided by large public displays, and if we like the things we see, we can download them to our mobile phone or at least save an URL to look at it in more detail later. But is the media we see really interesting to us? Is it worth being downloaded? There is a possibility to make the displayed information more interesting and useful: just let the people decide, which contents will be shown. And let the people influence what will show up next.

2.3.1 The MobiLenin System

Scheible and Ojala dealt with this subject in their MobiLenin project [10]. With the help of this application, users are able to choose out of a number of music videos. Short parts of six tracks - all the same length but different contents - are therefore shown on a large public display. The users have a mobile phone and are connected to the controlling server via HTTP over a GPRS data connection. Within a time interval of 25 seconds they get the chance to vote for one of the introduced video clips. They can do this with the help of a list showing up on the screen of their mobile phone (see figure 2). At the end of the voting time, the server counts the votes. Now the statistics of the poll and the winning track can be watched on the public display. As soon as the track is over, a new round starts introducing another six tracks. To encourage people to take part in the voting, one of the voters will be the winner of a lottery played by the server. He will receive a voucher for a beer or a pizza.



Fig. 2. MobiLenin: Screenshots of the client application [10]

The provided results of an user study show, that people are interested in such applications. Three cameras recorded the reactions and behavior of the participants, while a questionnaire asked all 14 of them about their experiences during the usage of the system. Obviously, the users liked the MobiLenin system: the camera recordings showed "laughing, happy faces, good mood, and rowdy celebrations upon winning in the lottery". 13 of the 14 users stated that the system was easy to use, because almost everybody is familiar with mobile phones. Comments like "The idea has a lot of potential" indicates once again that people are willing to interact with public displays and are able to imagine the usage in many different scenarios. It seems easy to imagine such a music video voting system in a subway station.

2.3.2 Jukola

A similar concept is implemented by Jukola, an interactive MP3 Jukebox [9]. Visitors of a bar can use provided handheld devices to choose a song out of four suggestions. By pressing additional buttons, they can also retrieve information about the artist, the album or similar bands. But it is not only possible to choose songs, people can also nominate them. Therefore a public display in form of a touch screen is installed in the public area of the bar. At any time, bar visitors can go up to the screen, browse the music collection and nominate songs by simply touching them. The third component of the system is a website, which provides the playlists of each day in the bar. This helps people to find out, what kind of music is played on a certain week day. Visitors can also recall the songs heard during last night's bar trip and thus remember the good time they had. The one week long evaluation of the system in a cafe bar was successful. The choice of a song became mainly a group event, rather than a decision by a single person. The groups started to talk and argue when a new choice was up. They tried to find a decision together and were waiting curiously for a new song to come up. The music choice even became competitive, and if a group voted for the song that was actually played next, they "seemed to feel a sense of pride that they were able to select the winner".

Unfortunately, it would rather be tough to install a system like this in our subway station. Touch screens are not suitable for public places like this, because they easily become subjects for vandalism and would thus create high expenses. Besides, the time before the next train will in most cases not be enough to vote for and listen to a full song.

Nevertheless, voting systems in general could inspire future research in the field of public displays. Representative political polls are one possibility, not only for a subway station: hundreds of people are potential participants. The providers of the public displays could also use it for little public-opinion polls to find out, how customers feel about certain aspects. The subway company for example could ask, if the waiting time for the next train is too long or not.

2.3.3 Polar Defence

Another possible way to interact with public displays is taking part in an interactive multi-user game. Polar Defence [6] is a very recent example. This game is designed to encourage a big number of people to anticipate in a group interaction with a large display on a public space. Unlike Shoot & Copy [2] or MobiLenin [10], Finke et al. tried to avoid any human interference while the players were being active. Instead of explaining everybody how to play, they installed a second large display showing information about how to participate and what to do. Besides text it used a huge animation which not only showed the rules of the game, but was also thought to attract bystanders from farther away. It could easily be seen from a distance of about 30 meters. The game consists of a virtual field comparable to a chess board having coordinates from A to I and from 1 to 9 (see figure 3). The players are supposed to place six towers on the field. When the game starts, enemies walk across the field. The defenders sitting on top of the placed towers automatically try to defend the field against those enemies. The more enemies are hit, the higher the score gets. The players could also watch a highscore list next to the playing field, which increased the competitive character of the game. After the towers are placed on the field, the actual game action lasts 40 seconds. Afterwards, a screen with the results of the current player is displayed.

The other way to avoid human intervention was the choice of the interaction medium: By using SMS, Finke et al. didn't need to hand out prepared mobile phones. Since every phone is capable of the Short Message Service, the players could use their own ones. But besides being widely used, SMS also brings some more advantages: it is a reliable service, each user can be identified by his number and the players can anonymously take part in the game. To place towers on the virtual field, it is only necessary to send a SMS containing six coordinates (one for each tower) to a number provided by the information display. The players took turns on a first-in-first-out base. This is altogether an easy-to-use and familiar way of interacting with a public display. The concept attracted a high number of activities during a four-day-long evaluation on the campus of the University of British Columbia. The team received 203 SMS from at least 57 different players. 37 of those users played at least twice. Finke et al. observed the public place constantly without letting people know who installed the displays. They noticed quite a number of positive results:

- the mentioned animation explaining how to play (2 meters x 1.8 meters) drew many peoples attention
- people communicated by explaining the rules to others or helping them to format the SMS
- group members celebrated success and tried to develop strategies together
- groups even started to take notes to optimize their strategies to get into the highscore list
- feedback messages showing the next player, game results or highscore lists seemed to trigger emotions



Fig. 3. Polar Defence: The virtual playing field [6]

The game Polar Defence was a great success. It was able to show once more, that people enjoy interactions with public displays. The application was not only able to entertain users, it could also manage to increase creativity and create communication. Nevertheless some users also criticized the use of SMS: their mobile phone contracts did not contain free text messages and they were not willing to pay for using the game. This fact is another problem on the way to interactive public displays. The hardware installation, software development and maintenance of the systems require an investment after all. Companies are in most cases not willing to pay money for products that will not pay off. Researchers should be aware of this. In the future they will be forced to think about business concepts which make public displays and the applications affordable.

2.4 Add

Additionally to being able to view and download contents of public displays, we can now also influence the contents provided. "Real" interactivity between users and displays can be recognized no later than now. But some of us might not be satisfied, yet. In the age of facebook, youtube, flickr and the like, people are more and more willing to share their media, opinions, not to say their lives. The technologies mentioned before (WLAN, SMS, etc.) don't only allow to download and influence contents, but also enable people to upload and share different types of media. But do we really want to add our personal documents to *public* displays? In the following sections some papers will be presented, which gave attention to this scenario.

2.4.1 A Video Commenting System

Honglu Du et al. [5] implemented a video commenting system. They had two main goals: First, they wanted to "make public displays interactive" and second, they want to "increase people's interaction with one another". Those intentions make this paper important on our way to interactive public displays. The setting for their experiment was a big classroom of an University in the USA (see figure 4). They chose a class with 254 students because of two reasons: students in big classes usually don't develop a strong group feeling. This was a good opportunity to find out if this particular feeling could be increased by the using the application. And the second reason was that the class would use the last week of the semester to show short videos that were developed during the last months. The video commenting system consists of three components: a client device, a server and a large public display. The client can be any device able to connect to the internet (possible are laptops or mobile phones). For the users it is simple to use the application: they have to go online, log in to a posting website and send a comment. All comments are sent to the server, which is also in control of the public display. Every four seconds, a new comment is viewed by the display. This happens on a first-in-first-out base. The large screen behind the teacher's desk shows the comments in a ticker below the video (which takes about 90% of the available area). A big advantage of the web-based system is the platform independency. It can easily be installed in any public display setting. Another new factor compared to the discussed ideas so far is the ability of the users to "speak out freely". This aspect opens up new ways of interaction, for example it is now possible to publish detailed opinions or propositions.



Fig. 4. The video commenting system in action [5]

The authors' "video commenting trial was successful". Of the attending 192 students 129 posted no less than 3115 comments. Considering that there were not enough client devices available and thus six to seven students had to share two computers, this number seems to be even more overwhelming. The evaluation of questionnaires approved

those impressions: 85.6% of the students had a strong interest in sending and reading the comments and 83.7% would like to use the system more often, just to name a few statistics. Also, the results showed an increase in the students' feeling of community. Furthermore, some good side effects appeared: some students were surprised how many different aspects of the movies they could find in the comments, that they had not thought about yet. Others picked up words they had never heard so far and looked them up in dictionaries later.

Field studies like this one often bring up new aspects that were not expected. In our process of going through the possible types of interaction between users and public displays, we are able to find and collect new positive aspects in almost every step. But we also discover problems we had not in mind so far. The just discussed user study is a good example: 19% of the committed comments were considered to be spam. Some of them were even offensive and "tended to provoke others". Considering the environment of an university classroom, the high number of spam messages was surprising to the research team. In their future work, they will try to "ensure quality" and plan to "add a moderating process on the server side". But overall they discovered that students were enjoying the activity, interacting and feeling a better sense of community.

2.4.2 Café Life in the Digital Age

A related research project was presented in 2006 by Churchill et al. [4]. In "Café Life in the Digital Age", they placed a "large-screen public, interactive community board" (called eyeCanvas) in a café in San Francisco (see figure 5). One goal of the installation was to present information: news, menus and other café-related contents, and art pieces of local artists. The first form of interaction offered was the possibility for guests to leave their email address in order to receive the café's newsletter. But the interesting aspect of the eyeCanvas was an artistic one: the touch screen provided an area called the "scribble".



Fig. 5. The eyeCanvas and examples of an artistic scribble and a comment [4]

With the help of a pen, people were able to leave little notes or to draw arbitrary pictures. At the bottom of the screen Churchill et al. placed a "carousel" for the created scribbles. User were able to navigate by moving the carousel to the left or right. According to the paper, people loved using these applications. They left 1466 messages or scribbles in only 2.5 months (for some examples see figure 5). Most people liked using the scribbles in an artistic way (about 50% of them were drawings). Other scribbles were mostly messages, either talking about upcoming events, the café or giving feedback about the eyeCanvas itself. Spam was no big problem in this case, "fewer than one 1% of scribbles were considered worthy of deletion". The authors see a reason for the small amount of spam messages: they believe that people did not want to be witnessed or watched leaving inappropriate messages and were afraid of possible social sanctions. But this feeling of

being watched also had another effect: only 1% of the people left personal contact information like email addresses or address details. For comparison: the café also has a suggestion book for his customers. 8% of the entries in this book contained contact details, so much more compared to the eyeCanvas. A privacy issue seems to be relevant when talking about interactive public display.

The end of the evaluation occurred after 14 months of the eyeCanvas' use in the café. The reason for this was that the touch screen finally stopped working. This incident shows two things: First, the project was extremely successful. After only four months, the team had already counted 392,164 touches. And second, touch screens are not the optimal input devices for interactive public displays. An unreachable display and mobile devices as input mediums are probably a better way in the means of financial matters.

2.5 Change

After being able to view, download, influence and add contents, the idea of an *interactive* public display becomes more and more precise. A fifth and final form of interaction seems to be the most difficult one regarding related research and experiments. The category "Change" describes the ability of users to change contents already being displayed on the public screen. Going through quite a number of papers, it seems that only a few research projects focused on this idea. Possible reasons will be discussed later on. First, some scenarios and ideas will be described, which could benefit from this kind of interaction.

Looking at the little pieces of art created by people using the eyeCanvas [4], we came up with the idea of an interactive art work. Imagine a large display on a public place. Interested people could be provided with a certain topic, alternatively they could create their own idea. In the next step, they are free to draw a little picture or write some keywords that come to their mind thinking about the given topic. Other people, maybe bypassers or spectators could possibly be inspired by the first drawings and are now allowed to add something to the screen or continue the picture drawn so far. In this way, a big piece of art could be created with not only one artist, but many participating persons. The painting could possibly tell many different stories about different people and their feelings.

Another idea was inspired by Carter et al. in their project "Digital Graffiti" [3]. In short, they used a Plasma Poster, a digital community poster board. People could for example add URLs to scientific articles. In a possible scenario for this application, they mentioned a person who highlighted a paragraph of a presented article, to draw the attention of others to it. Why not extend this good proposal to a whole new application? A display could present a news article and let users highlight words or word groups, adding short side notes or underline text passages. This would connect the topics described by the article to the opinion of the readers. Compared to the possibility of leaving comments, it would be easier to know, which part of the article the note is related to. Another possible scenario: students present a part of their homework (for example an essay) on a public display and let people correct wrong spelled words or propose suitable synonyms for wrong chosen nouns.

Thinking about these ideas, they raise one big problem, similar to the applications in the category "Add": spam. Giving people the opportunity to change something that already exists, there will be always some of them who think it's funny to destroy parts of it or to add offending material. It might be mainly this aspect, which hinders researchers to investigate this form of interaction.

3 COMBINING DIFFERENT CATEGORIES

So far we talked about projects that had one goal in common: making public displays interactive. But non of them exceeded the borders of their assigned categories. However there are researchers who tried to cross these borders and thus intended to create new opportunities. To take the good aspects out of different categories and combine them is a good way of collecting inspiration for new applications.

An outstanding example has been implemented by Izadi et al. Dynamo [7] is a system containing a large interactive display that can be used to "gather around, share, display and exchange media with

others." The basic scenario describes a group of people in a meeting, each of them having lightweight devices like USB drives, MP3 players and digital cameras, as well as more powerful ones such as PDAs or laptops. Sharing media files in such an environment with many participants can be complicated: exchanging memory cards or connecting laptops is rather time consuming and distracting. Right here the Dynamo application comes into action: providing "base interaction points" for lightweight devices and "mobile interaction points" for laptops and PDAs, group members can easily connect to the system. Each of them is part of a multi cursor environment, owning a pointing device identified by a certain color (see figure 6). They can now claim a carved region on the screen by holding down a mouse button and open up a rectangle of the desired size. Their connected device is shown in a "personal palette". They can browse the device and drag media files into their carved area (*Add*). Other participants can now look at the media file, for example a video or a document (*View*). If they are invited to share the carved area with it's creator, they can easily drag the items they like onto their own device and thus retrieve the media file from the partner (*Download*). A soon discovered problem was overlapping: two or more carved areas of different users got in each others way and made it difficult to move around certain files. Participants of evaluations solved the problem by informing each other and resizing their carved areas and moving around the media files on the screen (*Influence*). An additional feature of Dynamo is the users' ability to create notes. It is also possible to comment these notes and to edit them (*Change*).



Fig. 6. Dynamo: access, share, view, annotate and copy media files [7]

In several evaluations, users stated that Dynamo provided a simple way to exchange media files and that it is "much more sociable than using laptops around a table". But the user studies also uncovered some problems. Besides the overlapping problem, the participants needed instructions how and why they should create the carved areas. So a short introduction is needed before being able to use the system, but participants had no problems understanding the new interaction techniques. Some people expressed concerns, that others could simply steal files without permission. This problem was solved by letting only people download files from a carved area, who were invited by it's owner. Still, the privacy concern should not be ignored. Not only in these studies, but also when interacting with public displays in general, there will always be people watching. In a subway station for example, many people are present and a lot of them will focus on both, the person using the public display and the display itself. We can assume, that the users are aware of this fact and are not willing to share private information with people they do not know. In systems like Dynamo, with the help of avatars and carved areas, every viewer can identify, which files belongs to which user.

As you can see, Dynamo is able to combine features of all five introduced categories. But not only this fact is the reason for making it a powerful application. Used metaphors like drag & drop, cursors and icons make the application also understandable and easy to use.

Future works of the research team are supposed to connect remote Dynamo surfaces to enable remote and co-located workspaces. The idea satisfies the need for exchanging and showing media files to others, but has the problem of a strongly limited protection of privacy.

4 DISCUSSION

In this paper, we tried to follow a path from public displays with only minimalistic possibilities of interaction all the way to real interactive public displays. We discussed the process of transformation step by step: The first category, "View", represents the simple, information providing public display. No form of interaction besides viewing is given to the user. "Download", the second category, allows the viewer to retrieve interesting contents onto a mobile device. The process of remembering and reviewing information is now possible. By having an "Influence" on the contents of a public display, which is described by the third category, the user is able to chose what will show up next and at the same time express what he is not interested in. Fourth, people can "Add" their personal files, comments or drawings to the public displays and thus contribute to the contents shown and express their opinions and current feelings. And a final idea is to let users "Change" pieces of information that is already present on the screen. Besides this fifth category, all others were topics of very interesting and future-oriented research projects. Most user studies showed that people are interested interacting and communicating with public displays. Many of them had fun using the applications and wished that they could use them more often in the future. But despite of these successful evaluations, nearly all public displays found in everyday life are still not interactive. By taking a close look at all the papers, some possible reasons came up that might be able to answer this question.

First of all, it is a privacy issue. People are mainly used to devices like laptops, PCs, PDAs or mobile phones. Those are all *personal* devices, oftentimes storing sensitive data. Looking at a relatively small screen, people are aware of their privacy and they decide on their own who is allowed to see their information and who is not. Being faced with large public displays, many are concerned about their privacy. The interactive game Polar Defence [6] wanted to let players use their own mobile phones and used the numbers extracted from recieved SMS to identify each of them. The numbers were used to show who was up next or who made it to the highscore list. But they used only parts of that number, so that no spactator could identify any player. A major concern of Shoot & Copy [2] was also privacy: will people use the opportunity of downloading contents in public, although they might be watched by others? In this case, they were not, because they felt far enough away from the public display and thus it would not be possible for others to identify the contents they were interested in. The evaluation of the eyeCanvas [4] also showed the users' awareness of an affected privacy. Almost none of them was willing to leave any contact information or details about themselves. Standing in front of the large screen in the middle of the public area of a café they felt that people could watch them adding notes to the system. One of the first statements given during an interview with users of the Dynamo application [7] was the fear that provided files could be stolen by others without having permission. As a result, the researchers added a function letting only authorized people copy posted contents. As you can see, privacy is a major concern of people getting in contact with interactive public displays. We have to understand, that most of us are not used to the described forms of interaction being in public places. Future projects certainly need to consider this fact before their implementation. One finding is that the use of mobile phones is much more efficient regarding privacy issues than touch screens as input devices. Phone users can stay unidentified while everybody can see who is on the touch screen and sometimes even what he is doing.

Second it is money that plays an important role. It is certainly correct that the costs for the actual displays have been falling over the last years. But what we also have to think about are costs for additional hardware such as servers (in many cases controlling the communication between user and display) and provided mobile devices (like PDAs used for Jukola [9]). Operators of public displays also have to pay for the development and implementation of the applications and

a steady maintenance of hardware and software. And, living in the world of capitalism, we all know that companies don't make investments like this without seeing the chance of earnings in the end. The eyeCanvas [4] stayed part of the café for 14 months. Then, because of the high amount of users, the touch screen stopped working. The evaluation was extremely successful, but the owners were not willing to install the eyeCanvas permanently, anyway. Reasons were the lack of money for hardware and to pay a suitable administrator. One solution for the financing problem could be a fee for using the installed applications. But studies show that this would be hard to establish. Some players of Polar Defence [6] were not even willing to pay the cost for an SMS, which was necessary to take part in the game. So they found different ways: using a website providing free SMS, they simply used their laptops instead of mobile phones.

Third, it can sometimes be a problem to attract people to use applications on a public display. Since many of them are not used to scenarios like the ones described, they might be afraid of not being able to understand how to use the technologies. They might even fear to break something. Some efforts dealing with this problem have been made by the creators of Polar Defence [6]. In their experiment, they tried to attract many users without any human interference. They installed a second large screen, working as an information display: most of the available area was used by a huge animation, which could be seen from distances up to 30 meters. It was not only thought to attract bywalkers, but it was also supposed to explain the game rules with the help of six simple pictures. The rest of the screen was covered by some more information about the game and the number the SMS had to be sent to. The concept was successful, many participants played the game and nearly all of them were able to understand the rules without an explanation by the researchers. Hence, how to attract people to interact with public displays should be one of the major concerns of future projects.

And one more reason, why interactive public displays are not a constant part of our daily life until today, is the most annoying one: spam. Spam has been around ever since people were able to broadcast their lives. Even one of the most used communication medium of the digital age is endangered by spam: the email. Obviously, there will always be people who use new ways of communication to spread their commercials, offending comments or pointless statements. Just like everywhere else, spam has also been discovered in user studies of some of the introduced projects. The biggest spam problem occurred in a place, where the researchers expected it the least: the classroom of a respected university. 19% of all committed comments during an evaluation of the video commenting system by Du et al. [5] were classified as spam. In future work, the researchers find themselves forced to implement an additional module ensuring more quality of the committed comments. Ironically, a fact that had a negative effect on people's feelings about privacy, is now a possible solution to the spam problem: the eyeCanvas team [4] is sure, that people who think they might being watched post way less spam notes or offending pictures than people who are sure of their anonymity. Thinking about the category "Change", developers are probably aware that malicious users could always try to delete all contents or to destroy existing and carefully drawn peaces of art. It is a painful discovery, but only a few people are able to interfere with the effort and interest of many others.

5 OUTLOOK

We discussed both, positive and negative aspects of the interaction between public displays and their users. On the one hand, a great amount of work has been done. Researchers developed applications, which introduced new ways of interaction between the displays and their viewers. Userstudies showed, that many people have a vital interest in these kinds of interaction and enjoyed exploring them. Developers paid a lot of attention to the usability of the applications: using well known and intuitive technologies like mobile phones, mice or touch screens, they made it possible for users to find an easy entry in interacting with public displays.

On the other hand, researchers came across some serious problems. The most important one is privacy. Future projects need to be aware

that users are concerned about private data in public environments. If this problem will not be solved, people will most likely not adopt new possibilities of interaction. One of the solutions, assuring users of their anonymity, triggers another problem, which is spam. Developers need to focus on the quality of the contents of public displays and thus have to deal with this kind of unwanted data. Another issue that needs more attention is how to finance the installation of public displays and the connected costs. Obviously, users are not willing to pay for the applications. The next step would be to think about a business model that will encourage companies to invest in interactive public displays.

Future research projects dealing with public displays should not only keep up the work concerning usability and attractiveness of the applications, but also keep in mind the resulting problems. If they do, we are on a good way of making public displays interactive.

REFERENCES

- [1] *Oxford Student's Dictionary of English*. OXFORD University Press, 1 edition, 2001.
- [2] S. Boring, M. Altendorfer, G. Broll, O. Hilliges, and A. Butz. Shoot & copy: Using mobile phones for accessing information on large displays. In *Extended Abstracts of the Ninth International Conference on Ubiquitous Computing (Ubicomp)*, Sept. 2007.
- [3] S. Carter, E. Churchill, L. Denoue, J. Helfman, and L. Nelson. Digital graffiti: public annotation of multimedia content. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 1207–1210, New York, NY, USA, 2004. ACM.
- [4] E. F. Churchill, L. Nelson, and G. Hsieh. Café life in the digital age: augmenting information flow in a café-work-entertainment space. In *CHI '06: CHI '06 extended abstracts on Human factors in computing systems*, pages 123–128, New York, NY, USA, 2006. ACM.
- [5] H. Du, M. B. Rosson, J. M. Carroll, and C. Ganoe. "i felt more of a member of this class": increasing students' sense of community with video commenting. In *CHI EA '09: Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, pages 4405–4410, New York, NY, USA, 2009. ACM.
- [6] M. Finke, A. Tang, R. Leung, and M. Blackstock. Lessons learned: game design for large public displays. In *DIMEA '08: Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts*, pages 26–33, New York, NY, USA, 2008. ACM.
- [7] S. Izadi, H. Brignull, T. Rodden, Y. Rogers, and M. Underwood. Dynamo: a public interactive surface supporting the cooperative sharing and exchange of media. In *UIST '03: Proceedings of the 16th annual ACM symposium on User interface software and technology*, pages 159–168, New York, NY, USA, 2003. ACM.
- [8] J. F. McCarthy. Using public displays to create conversation opportunities. In *Workshop on Public, Community and Situated Displays at CSCW02*, 2002.
- [9] K. O'Hara, M. Lipson, M. Jansen, A. Unger, H. Jeffries, and P. Macer. Jukola: democratic music choice in a public space. In *DIS '04: Proceedings of the 5th conference on Designing interactive systems*, pages 145–154, New York, NY, USA, 2004. ACM.
- [10] J. Scheible and T. Ojala. Mobilenin: combining a multi-track music video, personal mobile phones and a public display into multi-user interactive entertainment. In *MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia*, pages 199–208, New York, NY, USA, 2005. ACM.