Ubiquitous Media for Collocated Interaction

[1]Giulio Jacucci, [1]Peter Peltonen, [1]Ann Morrison, [1]Antti Salovaara, [2]Esko Kurvinen, [1]Antti Oulasvirta

[1]Helsinki Institute for Information Technology (HIIT), Helsinki University of Technology TKK

[2]Elisa Oyj

Abstract Ubiquitous media – the digital, multimodal content available to users through ubiquitous computing devices – holds out the promise of contextualising media objects better by making them closer to and integrating them with real-world situations. In this chapter, we present two systems that exemplify the opportunities of ubiquitous media in collocated user interaction. CityWall is a public multitouch display that took the form of as a shop-front installation. MapLens is a mobile application that augments paper maps with digital content when viewed via a mobile screen. In user studies, we observed the importance of *public availability* in supporting users' engagement in the interpretation and utilisation of digital media. Two concepts in particular prove important: the notions of a *common stage* and *performative interaction*. Through a more extended analysis, we describe how these concepts are able to support the future design of ubiquitous media and provide challenges, particularly with regard to effortful configuration and immediate but superficial use.

Introduction

Has ubiquitous computing entered our lives as anticipated in the early '90s or at the turn of the millennium? In this last decade the processing of media combined with sensing and communication capabilities have been slowly entering our lives through powerful smartphones, multimodal game consoles, instrumented cars, and large displays pervading public spaces. However, the visionary formulations (Weiser 1991) and updated scenarios (Abowd and Mynatt 2000) have not been realised, despite the fact that the technology has become increasingly accessible.

Abowd and Mynatt (2000) identify three interaction themes for ubiquitous computing: natural interfaces, context-aware computing, and automated capture and access to live experiences. They also posit everyday computing as an area of development focusing on supporting the user with a continuously present interface, also addressing the periphery of the user's attention and connecting the physical and virtual worlds. While recent advances in mobile telephony and web tech-

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

nologies have been driven by social usage, how this could affect ubiquitous computing has not yet been defined. Previous views have considered social implications rather than making social use the target of the design (Abowd and Mynatt 2000), or focus on remote and mediated social interactions in mixed reality (Buxton 1997, Benford et al. 2005, Grudin 2002, Crabtree and Rodden 2008).

The ubiquitous user seems to be characterised as someone who needs situated and continuous assistance, access to media and the recording of experiences, or the provision of mediated interaction in mixed reality.

In our constructive and field-oriented research we have found it useful to characterise the ubiquitous user as an active user who opportunistically and skilfully uses technology rather than being assisted, a user for whom technology is a resource in constructing experiences rather than just using it to record them. Finally, and most importantly, our studies promote the view of the ubiquitous user as engaged in the social usage and sense-making of media, thereby integrating its use into their everyday lives (Jacucci et al. 2007a, Jacucci et al. 2007b, Salovaara et al. 2006).

In this spirit we want to enrich the understanding of the development of *ubiquitous media* systems: technologies providing users with digital multimodal content through ubiquitous computing devices. We want to explore new “archi-facts", where the technologies are embedded and become a part of the architecture of our urban environments. We also focus on materiality that tie together the physical artefacts and embodied interaction, and discuss how digital objects and interfaces become props in our social environment, rather than just media to be consumed.

Through our constructive and field-oriented work the contribution of this chapter is to outline aspects of the *public availability* of ubiquitous media and its relationship to collocated interaction. Robertson (2002) provides an account of how we begin to negotiate the public availability of artefacts and embodied actions. Through concepts such as the reversibility of perception (the body as an object in the world, that we and others perceive and also “the sentient body as it is lived by a particular person” (p. 308)) it is possible to explain how we can act and create meaning in an existing social and physical world. While this work points out the fundamental mechanisms of the availability of artefacts and actions in a social and physical world, implications that are relevant to the design of experiences in virtual spaces are also drawn.

Our focus is on the capacity of ubiquitous media to support embodied interaction, in particular via natural interfaces. The central features are how the “public availability” of ubiquitous media provides *common stages* and opportunities for *performative interaction*. Common stages are configured by users utilising features of ubiquitous media and provide a scene for social interaction. Performative interactions further indicate how media objects and interfaces are used as props in embodied and expressive acts.

We present two cases and report our observations from extensive user trials:

CityWall. In a study of interactions at a large multitouch display, we detail how the public availability of media supports embodied interaction and performative encounters. Here shared experience is constructed in embodied performances elicited by the interface and supported by the architectural configuration. Public availability is implemented by the multitouch interface, which encourages users to interact and by the media, which is produced out of mobile posting and web harvesting.

MapLens. We present how the physical-digital configurations of maps and media viewed in combination with the lens of a mobile phone support group interaction. Here navigation and coordination tasks are carried out collaboratively and the mixed media configurations provide common ground and place-making as important aspects of shared experiences.

Studies of Collocated Interactions in Public Spaces

User studies of collaboration around large public displays constitute a major part of the research on social interaction in collocated settings. In their study Russell et al. (2002) observed the benefits of visible physical actions (that facilitate learning from others), difficulties in developing clear turn-taking practices, and varying emerging ways to collaborate without anyone taking a leading role while participants were using a touch-screen display designed for small-group collaborative use.

Before users can start interacting with a public display, they have to withdraw from other activities they are engaged in. Hornecker et al. (2007) have noted the central role of access and entry points for publicly available technology. Entry points invite and entice people into engagement, provide an overview of the system, and draw observers into the activity. Access points are the characteristics that enable users to interact and join a group's activity. All these factors produce the *shareability* of the system, which refers to how a system engages a group of collocated users in shared interactions around the same content. Brignull and Rogers (2003) have suggested positioning public displays along traffic thoroughfares and describe the ways in which the interaction principles are communicated to bystanders.

Tangible interfaces are a key element for supporting face-to-face social interaction, as the physical interaction objects (such as touch-screens) eliminate the need for restricted and often singular input devices such as keyboards (Hornecker et al. 2006). Arias et al. (1997) also highlighted the importance of interacting with physical objects in a world that is becoming more and more virtual. Both aspects of reality – physical and virtual – have their strengths and weaknesses, and future systems should provide a flexible way to move in both worlds freely.

Collocated interaction has also been studied in the context of media sharing. Digital media archive studies on PCs have presented new collaborative methods for sharing media, such as the manipulation of content with gestures (e.g. Morris et al. 2006; Rogers et al. 2004; Wu and Balakrishnan 2003) and the visualisation of content for shared use (Shen et al. 2002). A study by Frohlich et al. (2002) provides a useful distinction between collocated media use practices, differentiating *storytelling*, mostly a single-person endeavour, from *reminiscing talk*, which is more a collaborative project, where many people participate, sharing their experiences of the same photograph with the others.

Studies on the collocated creation of media with mobile devices are, however, surprisingly few, probably because gathering data from this kind of use is difficult because of the personal nature of the devices. We organised a number of field trials at large-scale events such as a world championships rally, music festivals, and a national festival day analysing usage of group media applications through observations and systems logs (Jacucci et al. 2005a, 2005b, 2007a, 2007b; Salovaara et al. 2006). The studies highlight the aspect of *active spectatorship* (Jacucci et al. 2007a, 2007b) in participation in large-scale events. We observed users performing different expressive activities with the media: staging, arranging impromptu competitions, storytelling, joking, communicating their presence, and portraying others. With regard to remote others, the collocated users could engage in activities such as reporting what was going on onsite and making plans.

All these expressive forms contain creative elements such as joking, exaggeration, deception, and so on. The collocated use of ubiquitous media is thus a constructive activity, and not a passive endurance of events. Similarly, the active role of visitors to a museum has been noted by Heath et al. (2002), who studied how people construct the meaning of objects and artefacts together during museum visits. The creation and use of ubiquitous media in public is an eventful social activity, but the research is only starting to grasp its possibilities.

Another potential application area for ubiquitous media is tourism. Brown et al. (2003) presented an ethnographic study of city tourists' practices, describing how tourists work together in groups and collaborate around maps and guidebooks, which are used in combination to plan and create a setting providing an opportunity to spend time with friends or family. For this purpose the authors recommend systems that support sociality and, for example, combine a physical map with an electronic version of a guidebook.

CityWall: Multi-User Interaction at an Urban Multitouch Display

One of the visions of ubiquitous computing has been to render media accessible on large displays integrated into the environment. CityWall is a large multi-touch display installed in a central location in Helsinki, Finland. It acts as a collaborative and playful interface for the ever-changing media landscape of the city. The main features of the CityWall technology are: 1) multiple hand tracking capable of identifying uniquely as many fingers and hands as can fit onto the screen; 2) hand posture and gesture tracking; 3) high-resolution and high-frequency camera process-

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

ing up to 60 FPS, and 4) computer vision-based tracking that works in changing light conditions. The main motivation behind implementing CityWall in Helsinki was to support interactions for any user, from a child to a senior citizen, that did not require special skills or previous knowledge. The four technological features create the conditions for such a multi-user and multi-touch installation that is appropriate for a public space. Technologically, the setup is similar to HoloWall (Matsushita and Rekimoto, 2003). With this setup, all the expensive equipment can be placed indoors out of the public space, and after the addition of a semi-opaque thin coating, normal safety glass in a shop window can be used as a touch screen.

CityWall shows media content (pictures) about the events taking place in the city, photographed by Flickr users or by users sending them directly by email or phone. It gathers this content by querying pictures tagged with certain keywords ("Helsinki" and the selected large-scale event in our case). By this means, CityWall attempts to provide a sense of awareness to its users and the passers-by about both ongoing and past urban events, and a place for exploring these in a public site. Figure X.1 shows a screenshot from CityWall with Flickr content displayed on it. The bottom part (B) of the screen has a timeline with thumbnail-sized pictures. It is navigated by scrubbing it horizontally and it can also be compressed or expanded to show the contents retrieved during a full day or just a couple of minutes. This has been found to be important as the frequency of media may vary greatly.

Fig. X.1. Screenshot of CityWall with Flickr content.

Interaction with the top part (A) of CityWall follows two interaction paradigms. The moving, scaling, and rotation of content (C) follow direct manipulation principles: the user can grab an image by putting a hand on it. The photo fol-

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

lows the hand movements when the user shifts her hand. Rotation and scaling are possible by grabbing the photo at more than two points (e.g. with two hands or two fingers of the same hand) and then either rotating the two points around each other or altering the distance between them.

For a user study, CityWall was installed in a central location in Helsinki, Finland for the summer of 2007. The site was a 2.5-metre-wide shop window between the main bus and train stations (see Figure X.2).

Fig. X.2. CityWall installation in Helsinki, Finland.

The analysis presented below is an extension of our earlier findings (Peltonen et al. 2007, 2008), based on the analysis of eight days of use, during which 1199 persons interacted with the system in various social configurations. The data constituted webcam-quality video footage shot from a corner of the shop window's sunshade and interaction logs based on users' touch-based interactions. Selected parts of encounters were examined qualitatively. In this data, only 18% of the users were using CityWall alone. At other times, groups of users were present at the display, or people were watching others using the display and awaiting their turn. This made the usage a social experience.

While our previous work has focused on social interaction and turn-taking in general (Peltonen et al. 2008), in the following section we will specifically address the findings related to collocated interaction at the display.

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

Turn-Taking Management at the Display

To be able to browse the content together with others, people needed ways to manage turn-taking. Typically, the people who were present negotiated who had the right to use the display in a discrete manner. Sometimes this also led to game-like activities and competitions, especially between friends.

In multi-user situations, two basic settings exist. In *parallel use* people occupy an area of the screen and focus on their own task, irrespective of the activities around them. In doing this they consider the display estate in front of them as theirs, and assume that the others are doing the same in their part. In this way all the users have an equal chance to manipulate the content; however, they are not able to make use of the whole width of the display. Figure X.3 shows an extreme case of this, with seven users interacting with the display in parallel. In a few cases the users also engaged in teamwork, by teaming up to do the same task together, as when resizing the same photo together.

Fig. X.3. An extreme example of parallel use.

The other pattern of collocated interaction consisted of alternating *collaborative interaction*. For example, when the space in front of the display was crowded, members of a group (usually a pair) could organise themselves so that only one user used the display, while others contributed to the interaction by suggesting what to do next, being prepared to take the floor at transition relevant

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

places (Sacks et al. 1974), such as idle breaks or verbally negotiated moments. Figure X.4 provides a sequence that shows examples of both teamwork (a-b) and alternation (c-e). In Figure a, the girl in a white top is experimenting with resizing and rotation, assisted by her friend, who feeds her more photos from the right. In Figure b the girls are rotating one photo together, employing three hands in this task. In contrast, in Figures c-e the girl in the black top has the floor (c), but gives it to her friend (d), who then uses the display alone (e).

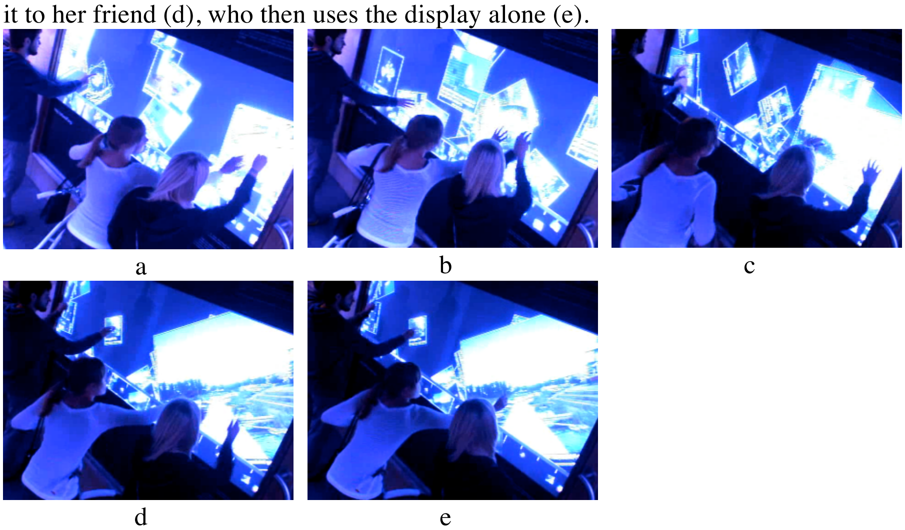

Fig. X.4. Teamwork (a-b) and alternating collaborative interaction (c-e).

In order to hold the floor, people did not only occupy the space in front of them. They could employ “pondering grips” (holding a picture steady and keeping their fingers on it) and “grandiose gestures” (large hand movements that act as an equivalent of shouting in verbal communication), to make clear that their interaction is still underway.

The activities of different groups are likely to collide at some point. We observed conflicts that mostly related to the ownership of photos and their immediate surroundings, i.e. areas that may be needed for rotating, scaling, and sorting the set of photos being worked on. Occasionally the UI causes people to break these territorial borders. For example, photos can be accidentally blown up or the timeline can be used without the other participants being considered. This was found to be the most disturbing conflict by the users interviewed.

Although conflicts take place, they can also have positive consequences for the social organisation at the display. Figure X.5 shows a sequence that starts with polite negotiation of space but turns into an impromptu game-like interaction between the parties present. The girl in a white top waits for her opportunity to interact with the display (a), and, after occupying a space, engages in alternating interaction with her friend to her right (b). After the girls have enlarged one photo to a very large size accidentally (c), the boy to their left manages the conflict by withdrawal (d), but later finds another space (e). A man, having observed the conflict and its resolution and been inspired by it, steps in and turns the conflict about the display estate into a game, first claiming one photo as his by shouting “It’s mine, don’t touch!” (f), after enlarging it to take up all the space in the display (g and h). In addition to screen estate competitions we also observed games such as Pong or soccer, which consisted of photos being thrown horizontally across the screen.

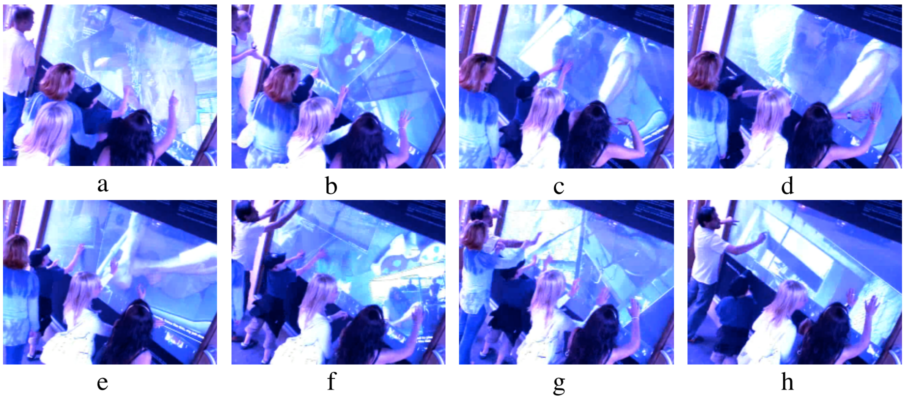

Fig. X.5. A sequence of negotiated interaction, conflict resolution, and game-playing.

The Display as a Stage

In the previous section we analysed the interaction between peers who assumed equal roles in interaction with the display. When people team up at the screen, individuals in groups may also gravitate towards complementary roles or social configurations. The most common social configuration of that kind was the teacher-apprentice setting (see Figure X.6). For instance, one or more users could take the role of an experienced user, and others attended to their presentation and demonstration of what CityWall was capable of doing.

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

Fig. X.6. A teacher-apprentice setting.

In other cases, a member of a group could use CityWall to draw attention from others in the group, or assume the role of a *comedian* to make the others laugh. The user could enlarge a funny picture to a large size on the screen, or demonstrate difficult-to-master manipulations to others. Role-taking was naturally important in game-like activities.

Such roles – teacher, apprentice, comedian etc. – can naturally change during the course of interaction. In the case of CityWall, this, however, happened rarely, because of the relatively short time that people spent at the display. For instance, a median duration of interaction for a pair of users was just 60 seconds in our data. In other cases, when users interact with different kinds of content, more role-switching may take place.

Shop windows and other places and objects in urban settings may enable or require people to interact with each other (e.g. when queuing or asking about an empty chair in a café). CityWall brought people together by making them interact collaboratively or with respect to each other in a space that was both digital and physical.

Ubiquitous technology in public space concretises Goffman's (1956) metaphor of social interaction as theatre. For instance, consider the situation in Figure X.7. The man in a white shirt enters the display (a), enlarges one of the photos to its maximum size, in this way occupying the whole screen and taking the floor from

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

the woman to his left (b). When rotating the photo, he shouts theatrically "the world is mine" (c) and leaves the screen laughing (d).

a b c d

Fig. X.7. Using CityWall as a stage for performance.

Conflicts between the parallel tasks of two or more users or teams were the main reason for the interactions between strangers that were observed. Users did try to avoid interfering with parallel activities, but the system did not support *the norm of social segregation* between the unacquainted, but made photos accidentally expand or fly across the screen. This forced the users to engage in conflict management with each other. The positive outcome was that the system made strangers interact with each other.

The photos displayed on CityWall were downloaded from public fora on the web. Therefore, the average user had no personal relationship with the content of the photos that happened to be on the screen when she appeared on site. The user's attention was thus turned from content to aspects of the interface. There were also users that seemed to take the content of the photos seriously, but the vast majority seemed to focus on playing with the interface.

MapLens: Place-Making with a Map-Based AR Application

Ubiquitous media can be interwoven with everyday environments through different types of natural interfaces (Abowd and Mynatt 2000). An example is the use of mobile devices for augmenting everyday objects, such as a map. MapLens is an application that allows users to read a standard non-marked paper map in tandem with a mobile phone. The application uses a *natural feature tracking method* (Wagner et al. 2008) to identify the map area visible beneath the phone's screen, augmenting the location with additional digital information displayed on the screen. Users click on icons to access larger versions of images or text (Figure X.8). Additionally, user-taken photographs are uploaded to the database, placed as per a set of GPS coordinates, and shared on-screen between all the users. The system allows for a distance of 15 to 40 cm between the printed map and the camera, and a tilt tolerance of up to 30 degrees. The paper map was printed onto an A3 size foam-core card, which was carried and used in tandem with the mobile device.

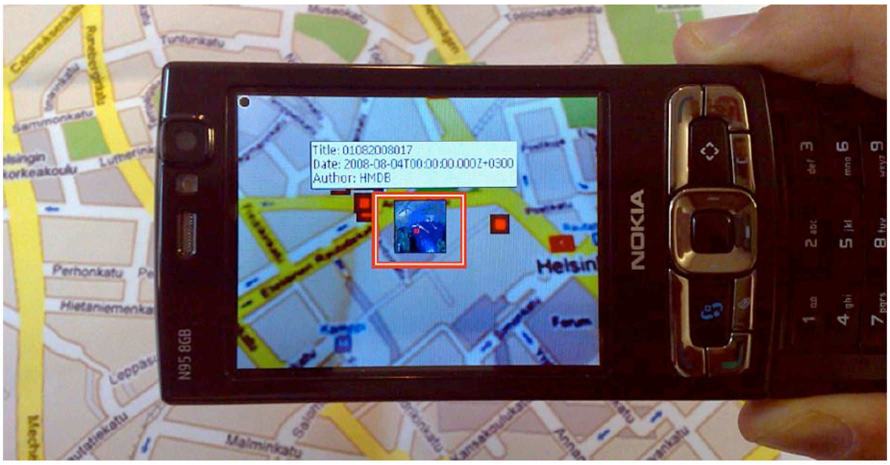

Fig. X.8. MapLens in use with a paper map, overlaying digital information on-screen. The red square (centre) is used to select augmented icons.

For the sake of comparison, we added a digital version, Google Maps mobile (Figure X.9). Both systems use the same virtual map and augmented information. To evaluate the technology, we devised an environmental awareness location-based game requiring players to complete twelve varied tasks (some sequential), negotiate roles, and coordinate the task order. The game was trialled over three Sundays in 2008 in the centre of Helsinki. Each trial was incrementally larger in size, with the final trial including DigiMap. We wanted to see if there were differences in how people used the two systems for the same tasks.

Fig. X.9. DigiMap version, Google Map with icons. Used by the control group in the trial.

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

We enlisted professionals, early adopters, environmental researchers, a scout group, and their friends and families. Thirty-seven people, 20 females and 17 males aged 7 to 50 years, participated. 21 had owned five or more mobile phones, with 22 familiar with the brand being trialled, and there was one non-phone user. Players were grouped into teams, and, using the technology, followed clues and completed given tasks within a 90-minute period. Each team managed the coordination of a kit of artefacts (see Figure X.10). One connected series of game tasks included: find a leaf in the museum; find the same leaf outside the museum; take a sunlight photo of the leaf using water to develop it (supplied in the kit; see Figure X.10); test pond water; test sea water for chlorine, alkalinity, and pH balance (supplied in the kit), and record all the readings (upload photos or enter the results in the clue booklet). To encourage friendly competitiveness we included three prizes, awarded for speed and accuracy, the best photograph, and the best environmental task design. The game promoted internal and external group activities: the negotiation of tasks and artefacts; 'noticing' and awareness of the environment; higher-level task management, and physicality, proximity, embodiment, and physical configurations around artefacts. There was particular emphasis on the mix of digital and augmented objects, with overtly tangible and tactile ones that were included to encourage physical proximity and team bonding and to 'jolt' users away from small-screen absorption.

Fig. X.10. Kits contained 7 items that needed to be managed: sunlight photographs, map, phone, water testing kits, voucher for internet use, clue booklet, and pen.

In our observations we noted the joint efforts of the users and present a discussion of our findings. In this section we label figures and name groups with M when referring to MapLens, and with D when referring to DigiMap.

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

The mobile device and bodily configurations

MapLens users (hereafter “M users”) typically held the device with their arms stretched out because the camera needed to be held at a range of 15-40 cm away from the paper map. The best light to view by was with sunlight on the map and the lens in shade. Importantly, with the device placed in this way, with one’s arm stretched out, others could see what part of the map was being examined, and, at times, the contents of the display.

In contrast, DigiMap users (hereafter “D users”) users typically held the device lower and closer to their body, as with a conventional phone. However, this rendered the phone more private (see Figure X.11 right) as others could not directly see the contents of the screen. Shading from the sun with the use of one hand, and, generally, one-handed use of the device was possible with D.

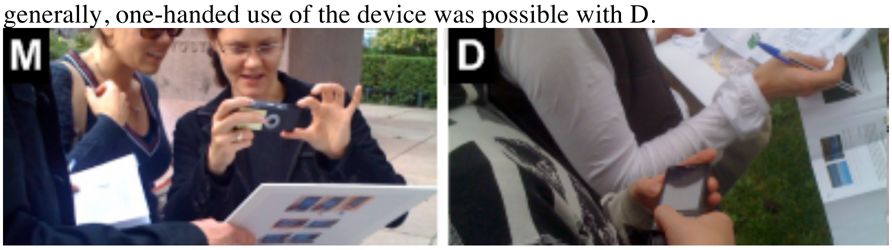

Fig. X.11. MapLens (M) was held in such a way that it could be shared with the group, whereas DigiMap (D) users held the device more privately.

Handing over of the phone occurred more with M groups than with D ones. When difficulties were encountered, it was common for the phone to be passed on. D user roles were defined earlier in the game, and one user tended to ‘own the use’ of the phone, while others managed parts of the kit.

Fig. X.12. Walking while using and bodily configurations. Left: Girls walk in front while one tries to read off MapLens (M). Center: MapLens (M) team negotiate where next. Right: One DigiMap (D) user reads the system while the other navigates.

We observed teams negotiating together in all parts of the trial. The discussions concerned not only the task at hand and what the team should do next, but also how best to use the technology (see Figure X.12 centre). In many instances M users gathered together around the physical map to use M. The group members who did not have the phone gave instructions to the one holding M about where to look. Needing to hold the map steady restricted movement (Figure X.12 centre),

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

unlike the case of D, where often one person was the ‘navigator’ of the group, searching for things from the mobile, while others observed the environment and led the way (Figure X.12 right). The bodily configuration around the use of D was separate and individual. The smaller screen and lower visibility meant that less sharing occurred and the division of roles took place earlier in the game.

Turn-Taking

Corresponding to the ease of handing over the phone, turn-taking, switching and transiting back and forth between roles spontaneously also occurred more with M. In Figure X.13 left we see Player A holding the map level, while Player B looks through the lens. Player C, close by, is listening and interjecting, while looking at the environment for clues. In Figure X.13 centre, Player B is still holding the phone, but the focus is on Player A and C looking at clues, reading instructions, and determining the location and the ordering of the clues. In Figure X.13 right, the team is posing barefoot on the grass (a game task). Player A is using the device as a camera, Player C is advising on camera use, and Player C is arranging their shoes in order to compose the shot optimally. The natural ease with which the passing of the phone and artefacts between the players occurred was reflected in the ease with which role-switching, turn-taking, and coordinating occurred.

Fig. X.13. Turn-taking: Left: PlayerA holds map as Player B looks through MapLens, Player C (out of picture) looks at environment. Centre: Player C looks at one clue while Player A looks at another and Player B holds device. Right: group poses for team photo barefoot on grass; Player A uses device to take photo, Player C points to device, Player B composes the shot.

Establishing Common Ground

Given that the typical way of using M involved a team gathering around and gesturing at the physical map with the device, establishing common ground was easier for M groups. We noted a shared understanding around the objects that were the focus of the co-conversants’ attention (Costabile et al. 2008). The location of M on the paper map, and the contents revealed to others on its display, helped everyone to understand the points under discussion without explicitly needing to

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

ing to ask. In Figure X.14 a young woman browses the map by using M. After finding a place, she suggests it to her father by pointing to it with her finger. The father proposes a nearby location and points to it by using the corner of a clue booklet.

Fig. X.14. The physical map as a common ground, established by showing with MapLens (M) and pointing with finger.

D teams were not able to share the map in such a fluent way (see Figure X.15). A young boy tries to identify a place by pointing to a relevant location on the screen and glancing around. After this he gestures towards the direction he suspects is correct, and hands the device over to his uncle.

Fig. X.15. DigiMap (D) Users experienced difficulties while attempting to share the map as common ground.

The physical paper map supported the players better in establishing a common understanding of the area and referring to different locations. The combination of MapLens and the paper map provided a means to be collaborative in a more physical way with other objects: fingers, clue booklets, pens, and other components from the kit (see Figure X.16 left). However, some M players found it challenging to identify the location on the map through the focus of the lens, especially while it was in use by another player. The D players often referred more directly by pointing at their surroundings.

For one D team we observed constant pointing at the mobile screen, establishing common ground. In another D team one looked at the screen behind the 'navigator's' shoulder (see Figure X.16 right), but with other D teams this did not occur.

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

Fig. X.16. Referring to objects by pinpointing. Left: pointing with a pen while using MapLens (M). Right: pointing with a finger from the DigiMap (D) screen.

Place-Making

Stopping, holding out MapLens and the paper map, and gathering around for a short time created an ephemeral opportunity, isolated from the surroundings by the physical map and the team members' bodies, to momentarily focus on a problem as a team. The phenomenon of place-making with the mobile use of technology has been raised previously in the literature (Kristoffersen and Ljungberg 1999) and here we encounter a special multi-user form of it. The physical map, as a tangible artefact, acted as a meeting point, a place where joint understandings could be more readily reached and participants were able to see, manipulate, demonstrate, and agree upon action. In pausing for discussion the teams created a series of temporary spaces, places for collaboration where they 'downed' bags, swapped or rearranged the objects they were carrying, stabilised the map and looked through M again to ascertain their progress. At this rapidly-made 'place', tasks were again shared, the negotiation and switching of roles, artefacts, and the device often occurred, and we noted a different kind of social usage in this temporary place, with other pedestrians walking around these 'places.'

Conversely, D teams only needed to stop at the places that the tasks themselves dictated; the rest of the action, decision-making, and way-finding was mainly done on the move or while stationary and completing tasks.

As a general overview, it becomes clear through observation, interviews, the results of the game, and observations that the MapLens users concentrated more on the interface, but not the environment around them. Additionally, the MapLens users were more concentrated on the combination of the technology and the game – which involved problem-solving via negotiation and physical and social interaction. The way in which place-making affects attention on the task and technology, as opposed to the surroundings, is a plausible explanation for this observation.

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

Discussion

The cases presented our understanding of ubiquitous media by providing observations from real-world use. Nearly a decade ago Abowd and Mynatt (2000) recognised that the evaluation of ubicomp systems is difficult. Since then, numerous proofs-of-concepts have been presented, but empirical studies in real settings are rare. Neither have conceptual work for tangible interaction (e.g. Hornecker & Buur 2006) or shareability (Hornecker et al. 2007) been systematically informed more by anecdotal evidence than by empirical, prototype-driven field trials (for an example, see Huang et al. 2007).

While most of the previous work on ubiquitous media has focused on distributed and mediated interactions, this chapter looked at collocation. We presented two examples: in the CityWall study usage included mostly playful interaction and engagement with others, while users were less engaged with the content (pictures and text concerning the city). In the MapLens case the interaction was more collaborative; the users focused more on the content and interacting with each other. The trial included a control condition where a non-augmented version (DigiMap) was utilised. Although the DigiMap users were faster in completing their tasks, the MapLens users were more thorough in their execution and more engaged in the game, and their solution process was more collaborative.

Behind the two designs, we can observe two concepts of how ubiquitous media renders media publicly available for others. We conclude the paper with discussion of these concepts, with the purpose of informing future design.

Common Stages

In the MapLens case, collocated interaction took place mostly while on the move. Mobility is characterised by the uncertainty of events and courses of action (Perry et al. 2001). At its best, ubiquitous media can provide open and flexible support for groups that, in the face of unexpected events, need to coordinate their joint activities. Maps and guidebooks can provide this support, as was found by Brown et al. (2003). The key question for designers is how technology comes in between people, segregating or connecting them in ways inconsistent with the prevalent orderliness of social action in urban space.

Ubiquitous computing technology provides novel means to establish and become drawn into *face engagement*, where a sense of *mutual activity* is formed (Goffman 1963). Furthermore, after being drawn in, since it is largely unclear to people how to behave in this new situation, people rely on each other to learn about the possibilities and constraints of the setting. The metaphor of a *stage* (Goffman 1956) is the most descriptive of the ways in which the public availability of media in a physical space can create areas inside which the actions of individuals turn into a *performance*, with people assigning *roles* to themselves, as well as to others, and swapping them on the fly.

In both cases we observed how the ubiquitous media contributed to the formation of a common stage for collocated interaction.

In the case of CityWall, common stages were facilitated by the public availability of the media, with passers-by being able to easily learn how to use the stage by watching others interacting with the display and with each other. The stage was framed by the architectural solution, the large size of the display and the space in front of it. More importantly, the stage was framed by people approaching, negotiating, and using the display. On this common stage people could adopt roles such as being teachers, apprentices, clowns, or members of the audience. In some cases, multiple activities were taking place at the same time. Such asymmetric participation patterns have not been reported to this extent before. This stage provided an opportunity for encountering 'the other' (Schutz 1967) with the possibilities of being in possession or control of the ubiquitous media and its interactivity. In particular, the single timeline provided a strong feature for encountering others through coordination or conflict.

In the MapLens study, the map and mobile device as tangible objects provided a common ground to perform fine gestural communication movements allowing reference points during collocated interaction. This contrasted with DigiMap, where the tasks were carried out with more divided roles and without a common support as a reference. The awkwardness of the map and mobile device also led the users to place-making. The map and mobile device became the place around which to gather, acting as a pretext to appropriate a bench or a space in the park. The paper-card map was also used as if it was a miniature stage. All in all, the casual approach to the environment, the 'throwing-down' of kits and artefacts, and 'owning the space', supported rapid place-making. The kit of lo-fi tangible objects – a muslin bag, expandable paper clue booklet, biro, and a map on foam-core card – contributed to the informality of the use (Ehn and Kyng 1992). We observed how interactions and bodily configurations made the teams unaware of their surroundings and others in the environment, e.g. other pedestrians needing to walk around their rapidly made 'place', which could also be viewed – in Goffman's sense of the word – as an informally established stage.

Performative Interaction

An important aspect of collocated interaction is the shared experiencing of media. Heath et al. (2002) observed how the meaning of works of art was reflexively constituted in collaboration by museum visitors. Experiences are lived through, constructed, and interpreted as a group rather than as an individual (Forlizzi and Battarbee, 2004). This aspect was particularly accentuated in the CityWall examples

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

through bodily expressions and role-taking, causing the ubiquitous media to take on the role of a “prop” in *performative interactions*.

Here we refer to the idea of “interaction as performance” (Jacucci 2004), originating from pragmatic and anthropological views of experience and performance. For these views performance implies at the same time experience and expression, action and representation, consciousness of the act, and, like an event, an initiation and consummation.

For the MapLens users, holding and looking though the lens and ‘showing’ the other group members possible solutions on the paper card map forced a form of performative interaction where the phone became a physical interface and prop; it became an AR lens used to negotiate the floor, orient attention, and access information collaboratively.

Crabtree et al. (2005) argue that technology-assisted playful activities rely on the participants’ ability to manage, diagnose, and repair interruptions. Conversely, we have shown how playful and jocular activities help people not only to account for mistakes they make, but also deal with system-generated problems. In consequence, the collaborative learning process becomes more enjoyable and hence more sustainable over time.

In CityWall, the content on the wall and the features of the interface were used as resources to coordinate the activity and to create events or interactions, so they were meaningful in front of others. For example, interactions such as photo-moving and scaling turned into games such as playing Pong. Some gestures were made salient to others. “Grandiose gestures” and “pondering grips” were used to manifest the volume of the user’s actions and her intentions towards others, while also marking the boundaries of the workspace that the user felt she had claimed as her own. The presence of strangers – all the other people walking past the installation, sometimes stopping by to observe what the users are doing – also has an effect on one’s activities at CityWall, which can be perceived as a public performance in a city space.

It is clear that the use of CityWall provided richer instances of performative interaction than that of MapLens. This is also because of the explorative and casual setting of CityWall and because the use of MapLens took place in a game with goals and tasks.

It is not only about the interface being *big enough* to be used in parallel by several people (the case of CityWall). An impromptu stage for performative interaction can also be formed around smaller items, such as smartphones or printed maps (the case of MapLens). Paper, just like public displays, can also configure spaces and surfaces (Rogers et al. 2004) in support of social interaction. Both of our prototypes were able to support expressive gestures that helped the participants in coordinating, communicating, and acting out different roles.

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

Interaction Design Issues

To conclude, we would like to reflect on two design issues. Addressing these in the future would strengthen collocated interaction as an aspect of ubiquitous media.

Effort and common stages. The size and other aspects of the materiality of ubiquitous media facilitate or hamper collocated interaction. Surprising aspects, such as architecture and positioning in a physical space, affect use. Similarly, combining an artefact such as a map with dynamic media provided opportunities for social interaction. In the MapLens case effort was needed for technical and ergonomic reasons, such as illumination issues constraining the orientation of the map and the orientation of the device to the map, or the position of participants so that they could view and access the interface. This effort and awkwardness provided a pretext for place-making and common stages.

The approach of natural interfaces, such as multitouch on the one hand and augmenting physical objects on the other, also provided parallel access to media and shared experience. Challenges included the effort needed by the users to manage this availability. As an example, common features, such as the timeline in CityWall or the map and the mobile device in MapLens, need to be negotiated through coordination and conflict, which provides opportunities for shared encounters. The challenges are therefore to solve the trade-off between effortless operation and the sociality of the system.

Performative interaction for immediate but superficial use. In the case studies ubiquitous media was successful in collocated interaction in the way it provided props for interaction. We see the mobile device in MapLens used as a physical lens, or the media objects in CityWall used for playful and expressive interactions. Embodied and expressive interactions were easy to learn, playful, and immediate. At the same time there was the problem that the encounters relied heavily on the immediacy of the interface and its playfulness. The success of the usage, for example, of CityWall seems to rely on the novelty of the technology. Moreover, the immediacy and playfulness of the system attracted more attention than the media that were provided.

Acknowledgements

We thank Tommi Ilmonen and John Evans, who worked on the CityWall. MapLens and DigiMap were developed in cooperation with the University of Oulu, Graz University of Technology, and the University of Cambridge, within the IPCity project funded by the 6th Framework Research Programme of the EU. We thank Saija Lemmelä and Jaana Juntunen for their contribution in the analysis of MapLens trials.

References

1. Abowd, G. D. & Mynatt, E. D. (2000). Charting past, present, and future research in ubiquitous computing. ACM *Transactions on Computer-Human Interaction*, 7(1), 29–58.
2. Arias, E., Eden, H., and Fisher, G. (1997). Enhancing communication, facilitating shared understanding, and creating better artifacts by integrating physical and computational media for design. In *Proceedings of the 2nd Conference on Designing interactive Systems: Processes, Practices, Methods, and Techniques* (DIS'97, pp. 1–12). New York: ACM Press.
3. Benford, S., Magerkurth, C. & Ljungstrand, P. (2005). Bridging the physical and digital in pervasive gaming. *Communications of the ACM*, 48(3), 54–57.
4. Brignull, H. & Rogers, Y. (2003). Enticing people to interact with large public displays in public spaces. In *Proceedings of the IFIP International Conference on Human-Computer Interaction* (INTERACT 2003, pp. 17–24). Amsterdam: IOS Press.
5. Brown, B. & Chalmers, M. (2003). Tourism and mobile technology. In *Proceedings of the Eighth Conference on European Conference on Computer Supported Cooperative Work* (ECSCW 2003, pp. 335–355). Norwell, MA: Kluwer Academic Publishers.
6. Buxton, W. (1997). Living in augmented reality: Ubiquitous media and reactive environments. In K. Finn, A. Sellen & S. Wilber (Eds.). *Video Mediated Communication* (pp. 363–384). Hillsdale, NJ: Lawrence Erlbaum Associates.
7. Costabile, M. F., De Angeli, A., Lanzilotti, R., Ardito, C., Buono, P., & Pederson, T. (2008). Explore! Possibilities and challenges of mobile learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI'08, pp. 145–154). New York: ACM Press.
8. Crabtree, A., Rodden, R. & Benford, S. (2005). Moving with the times: IT research and the boundaries of CSCW. *Computer Supported Cooperative Work*, 14, 217–251.
9. Crabtree, A. & Rodden, T. (2008). Hybrid ecologies: understanding cooperative interaction in emerging physical-digital environments. *Personal and Ubiquitous Computing*, 12(7), 481–493.
10. Ehn, P. & Kyng, M. (1992) Cardboard computers: mocking-it-up or hands-on the future. In J. Greenbaum and M. Kyng (Eds.) *Design at work: cooperative design of computer systems*. (pp. 169–196) Hillsdale, NJ: Lawrence Erlbaum Associates.
11. Forlizzi, J. & Battarbee, K. (2004). Understanding experience in interactive systems. In *Proceedings of the 5th Conference on Designing interactive Systems: Processes, Practices, Methods, and Techniques* (DIS'04, 261–268). New York: ACM Press.

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

12. Frohlich, D., Kuchinsky, A., Pering, C., Don, A., and Ariss, S. (2002). Requirements for photoware. In *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work* (CSCW'02, pp. 166–175). New York: ACM Press.
13. Goffman, E. (1956). *The Presentation of Self in Everyday Life*. New York: Anchor Books.
14. Goffman, E. (1963) *Behavior in Public Places*. New York: Free Press.
15. Grudin, J. (2002). Group dynamics and ubiquitous computing. *Communications of the ACM*, 45(12), 74–78.
16. Heath, C., Luff, P. K., vom Lehn, D., Hindmarsh, J. & Cleverly, J. (2002). Crafting participation: designing ecologies, configuring experience. *Visual Communication*, 1(1), 9–34.
17. Hornecker, E. & Buur, J. (2006) Getting a grip on tangible interaction: a framework on physical space and social interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI'06, 437–446). New York: ACM Press.
18. Hornecker, E. Marshall, P. & Rogers, Y. (2007). From entry to access - how shareability comes about. In *Proceedings of the 2007 Conference on Designing Pleasurable Products and interfaces* (DIS'07, pp. 328–342). New York: ACM Press.
19. Huang, E.M., Mynatt, E.D. & Trimble, J.P. (2004). Displays in the wild: understanding the dynamics and evolution of a display ecology. In *Proceedings of the 4th International Conference on Pervasive Computing* (Pervasive 2004, 321–336). Berlin: Springer.
20. Jacucci, G. (2004). *Interaction as Performance. Cases of configuring physical interfaces in mixed media*. Doctoral Thesis, University of Oulu. Oulu: Acta Universitatis.
21. Jacucci, G. & Salovaara, A. (2005a). Mobile media sharing in large-scale events - beyond MMS. *Interactions*, 12(6), 32–35.
22. Jacucci, G, Oulasvirta, A., Salovaara, A. & Sarvas, R. (2005b). Supporting the shared experience of spectators through mobile group media. In *Proceedings of the 2005 international ACM SIGGROUP Conference on Supporting Group Work* (Group 2005, 207–216). New York: ACM Press.
23. Jacucci, G., Oulasvirta, A. & Salovaara, A. (2007a). Active construction of experience through mobile media: a field study with implications for recording and sharing. *Personal and Ubiquitous Computing*, 11(4), 215–234.
24. Jacucci, G., Oulasvirta, A., Ilmonen, T., Evans, J. & Salovaara, A. (2007b). CoMedia: mobile group media for active spectatorship. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI'07, pp. 1273–1282). New York: ACM Press.

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

25. Kristoffersen, S. & Ljungberg, F. L. (1999) "Making place" to make IT work: empirical explorations of HCI for mobile CSCW. In *Proceedings of the international ACM SIGGROUP Conference on Supporting Group Work* (SIGGROUP 1999, 276–285). New York: ACM Press.
26. Matsushita, N. & Rekimoto, J. (2003). HoloWall: designing a finger, hand, body, and object sensitive wall. In *Proceedings of the 10th Annual ACM Symposium on User interface Software and Technology* (UIST 2003, 159–168). New York: ACM Press.
27. Morris, M.R., Huang, A., Paepcke, A. & Winograd, T. (2006). Cooperative gestures: multi-user gestural interactions for colocated groupware. In *Procceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI'06, 1201–1210). New York: ACM Press.
28. Peltonen, P., Salovaara, A., Jacucci, G., Ilmonen, T., Ardito, C., Saarikko, P. & Batra, V. (2007). Extending large-scale event participation with user-created mobile media on a public display. In *Proceedings of the 6th international Conference on Mobile and Ubiquitous Multimedia* (MUM'07, pp. 131–138). New York: ACM Press.
29. Peltonen, P., Kurvinen, E., Salovaara, A., Jacucci, G., Ilmonen, T., Evans, J., Oulasvirta, A., & Saarikko, P. (2008). "It's mine, don't touch!": interactions at a large multi-touch display in a city centre. In *Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems* (CHI'08. 1285–1294). New York: ACM Press.
30. Perry, M., O'Hara, K., Sellen, A., Brown, B. & Harper, R. (2001). Dealing with mobility: understanding access anytime, anywhere. *ACM Transactions on Computer-Human Interaction* (TOCHI), 8(4), 323–347.
31. Robertson, T. (2002). The public availability of actions and artefacts. *Computer Supported Cooperative Work*, 11(3–4), 299–316.
32. Rogers, Y., Hazlewood, W., Blevis, E. & Lim, Y.-K. (2004). Finger talk: collaborative decision-making using talk and fingertip interaction around a tabletop display. . *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI'04, 1271–1274). New York: ACM Press.
33. Russell, D.M., Drews, D. & Sue, A. (2002). Social aspects of using large public interactive displays for collaboration. In *Proceedings of the 4th international Conference on Ubiquitous Computing* (UbiComp 2002, 229–236). London: Springer.
34. Sacks, H., Schegloff, E.A. & Jefferson, G. (1974). A simplest systematics for the organization of turn taking in conversation. *Language*, 50(4), 696–735.
35. Salovaara, A., Jacucci, G., Oulasvirta, A., Kanerva, P., Kurvinen, E. & Tiitta, S. (2006). Collective creation and sense-making of mobile media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI'06, 1211–1220). New York: ACM Press.

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.

36. Schutz, A. (1967). *The Phenomenology of the Social World*. Chicago: Northwestern University Press.

37. Shen, C., Lesh, N., Vernier, F., Forlines, C. & Frost, J. (2002) Sharing and building digital group histories. In *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work* (CSCW 2002, 324–333). New York: ACM Press.

38. Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T. & Schmalstieg, D. (2008) Pose tracking from natural features on mobile phones. In *Proceedings of IEEE International Symposium on Mixed and Augmented Reality* (ISMAR 2008, pp. 125–134). IEEE Press.

39. Weiser, M. (1991). The computer for the 21st century. *Scientific American*, 265(3), 94 –104.

40. Wu, M., Balakrishnan, R. (2003). Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proceedings of the 16th Annual ACM Symposium on User interface Software and Technology* (UIST'03, 193–202). New York: ACM Press.

In: Willis, K.S.; Roussos, G.; Chorianopoulos, K.; Struppek, M. (Eds.): Shared Encounters, Springer, 2009.