Configuring Gestures as Expressive Interactions to Navigate Multimedia Recordings from Visits on Multiple Projections

Giulio Jacucci [1] [2]
[1] Department of Information Processing Science,
University of Oulu,
P.O. Box 3000, 90014 Oulu, Finland
[2] Helsinki Institute for Information Technology
(ARU), P.O. Box 9800, FIN-02015 HUT

Juha Kela, Johan Plomp
VTT Electronics,
Kaitoväylä 1, P.O. Box 1100,
90571 Oulu, Finland
juha.kela@vtt.fi, johan.plomp@vtt.fi

ABSTRACT

The wide availability of digital recording devices leads to investigate how multimedia content can be navigated beyond a desktop computer set up. We present a system and a variety of applications, to navigate multimedia recordings from visits making use of: large multiple projections, location information to organise media and re-experience aspects of visits, physical interfaces as gesture-based interaction and other interfaces that render media more tangible and therefore more readily available. Analysing field trials we discuss the *expressiveness* and *experiential aspects* of gesture based interfaces as important features in navigating multimedia in "immersive" environments.

Author Keywords

Gesture-based interaction, visualising digital photographs, multiple projections.

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Moving interaction from the virtuality of the screen of a desktop computer to the physical environment can result in diverse challenges and opportunities, according to the activity that computers seek to support. A particularly interesting activity to investigate for ubiquitous computing is digital photography. The widespread use of digital cameras is resulting, in leisure and work settings, in using the computer (also television screens, and home theatre technologies) to view and share pictures often taken at

remote sites or during visits. This paper seeks to explore opportunities and challenges of providing novel physical interfaces to visualise pictures and sounds from visits, moving the viewing and sharing of pictures from a desktop computer to a ubiquitous setting. In particular, we investigate three opportunities. Firstly, large multiple projections can provide more immersive environments to visualise digital media and sound can play a more important role. Secondly, location information can be used to organise media and re-experience aspects of visits. Finally, physical interfaces, as gesture-based interaction, can free users from the desktop and other interfaces can render media more tangible and therefore more readily available. A variety of questions and challenges accompany these opportunities. This paper explores novel interfaces to navigate collaboratively pictures from visits. In particular, we discuss the *expressiveness* and *experiential aspects* of gesture based interfaces as important features in navigating multimedia in "immersive" environments.

For this purpose, we use a distributed environment developed in the Atelier project[1] that provides: computational support to record location information of recorded multimedia during visits, a hypermedia database to store visits as *hyperdocuments*, and a variety of physical interfaces to navigate the hyperdocuments. Physical interfaces include: barcode scanner and print-outs with thumbnails and barcodes, an infrared remote control interface, and a gesture based interface. We organised trials, where pair of participants, after having recorded a visit, navigated the multimedia recordings using two back projection screens. The contribution of this paper is twofold: to present a variety of applications to support gesture based-interaction in navigating multimedia recordings from visits on multiple screens, and to provide insights from user trials on the advantages and challenges of gestures as an interaction mode in such environments.

---

[1] http://atelier.k3.mah.se

The paper, first, briefly discusses related work in gesture-based interaction and novel interfaces for digital photography. We then discuss in detail the research questions and methodology. After that, we describe the Atelier environment clarifying which tools make possible to configure the navigation of multimedia recordings. A section is devoted to presenting highlights from the trials. The discussion summarises the lesson learned in constructing a more immersive environment to navigate multimedia recordings from visits, in particular addressing configurability and the advantages of gesture-based interaction.

Related Work

One of the ways that embodied actions or gestures are related to physical interfaces is movement. Benford et al [1] provide a framework centred on the notion of movement for physical interfaces: "new 'physical interfaces' considerably extend the repertoire of available movements to include whole body gestures, moving objects across surfaces, and moving the entire interface through space. Their aim in so doing is to make interaction "more natural, expressive, immersive or ubiquitous." Benford et al. [1]. Bodily movements have been considered to position or track a person's body, to track the movement of an object in the hands of a person, or to recognise gestures and facial expressions. Technologies to track and locate persons have been used to develop guides [2] and other location-aware information systems. Some aspects of nonverbal communication, e.g. visual expression (facial expressions, physical appearance, direction of gaze, physical posture), have been applied mostly to interact with virtual characters or for video conferencing. Gesture recognition systems, on the other hand, have been used to recognise sign language, for multimedia information kiosks, to control desktop applications and to interact with in-car devices [3]. Bodily expressions and interactive technologies have been the object of study in artistic performance [4]. Otherwise the expressiveness of gestures has been less the focus of research, which generally aims at making interactions more intuitive or usable and with a focus on single user interactions, e.g in Bellotti et al. [5] "develop systems that can communicate more naturally and effectively with people" (p. 423). For example, Swindells et al. [6] present a system for identifying devices through a pointing gesture using custom tags and a custom stylus called the gesturePen. Similar pointing technology can also be integrated into different mobile devices as proposed by Ailisto et al. [7]. Sparacino et al. [8] developed "a "media actors" software architecture used in conjunction with real-time computer-vision-based body tracking and gesture recognition techniques to choreograph digital media together with human performers or museum visitors." Sparacino et al. apply this to dance, theater, and the circus, augmenting the "traditional performance stage with images, video, music, and text, and are able to respond to movement and gesture in believable, aesthetical, and expressive

manners." Expression in interaction has been envisioned in what McGee [9] calls Contact-expressive devices, "…technologies that understand and use touch in meaningful ways—that can distinguish between a press and a caress." Most of the work in this area is oriented at developing the technology, while studies in real settings are hard to find. Also much work in tracking bodily movement has been applied to mixed reality environment supporting remote collaboration (Luff et al. [10]), which is not within the scope of this work.

Novel interfaces for digital photography have been concentrating on information appliances more than in multiple large screens. Balabanovic et al [11] describe an easy-to-use device, "StoryTrack", that enables digital photos to be used in a manner similar to print photos for sharing personal stories. Other work focused on sharing and sending photographs through mobile devices [13]. Büscher et al. [12], in the WorkSPACE project, addressed with a comprehensive approach and technology the support for flexibility, mobility, and collaboration for landscape architects. They present one mobile appliance, the "sitepack", and two room appliances (a panel with pen interaction and a table with a horizontal display area). All three appliances run the Topos software in order to share workspaces and documents. The sitepack provides ways to index photographs and sound notes with GPS coordinates, and is also designed as a remote collaboration tool. The large display is rather used to dispose documents in a virtual environment.

Research Approach

Elsewhere, we have reported on how the development of the environment described in this paper, has been informed by field studies and trials on architecture students visiting remote sites [14, 15]. In this paper, we present more details on the configuration tools, new technological developments and findings from a new trial.

1) In previous trials we used two projection screens to display on one the pictures and on the other a 3D visualisation of the current position on a map. Participants could move forward and backward in the virtual 3D map between positions (or nodes) that contained media recordings.

2) In the new trails, both screens are used two visualise multimedia recordings or 2D maps according to user's preference. In particular, thumbnails of multimedia recordings and layouts of 2D maps can be printed out with barcodes, to provide a physical interface to the digital content.

3) We have implemented configurable interfaces to select the current multimedia output (separate media player on each screen or different sound outputs).

4) The mobile application had also support to store the compass direction of the taken picture. Now, this compass information can be used in the environment to navigate

through a group of pictures that were taken in the same spot (forming for example a 360 degree panoramic view) by pointing in different directions.

5) The new trials focused more systematically on the use of gesture-based interaction in collaborative viewing of recordings of visits.

The approach adopted in the trials, included asking two participants to choose a place of interest. They would then visit this place, using Atelier equipment taking multimedia recordings while performing paths, for example, walking. The participants were then invited in the Atelier environment, to navigate collaboratively the recorded "multimedia path", and the trials ended with an interview. The scenario they were asked to perform, was therefore clearly defined, but allowed participant to choose a place, choose what to record, and also how to navigate it together. The trials were conducted as explorative and formative evaluations, trying to let users contribute with their own ideas how to use the technology. We did not evaluate particular usability aspects or measure specific performances. We were more interested in gathering insights on the appropriateness and opportunities of gesture-based interaction to navigate multimedia recording from visits on multiple media outputs.

AN ENVIRONMENT TO NAVIGATE VISITS

The use scenario we set out to support is of one or more persons visiting a remote site of interest, where they take pictures and record sounds. In addition, the system automatically records the path according to the user's movement (Figure 1). At a later time this "multimedia path" can be collaboratively navigated using multiple screens, gestures and printouts with thumbnails of the recordings. The environment is therefore described in two steps. First, the recording of the visit using a mobile application and devices is explained, also describing how multimedia paths (HyperDocuments) are stored in the Atelier environment. Second, we describe configuration tools to set up the environment for navigation. After the two steps, we present insights from the trials.

Figure 1. Performing a visit recording sounds, and pictures

Performing a visit, recording media

The path is recorded using timestamps and GPS (Global Positioning System) trace. These are created by a mobile application (eDiary), while visitors take pictures, videos, sounds, and text notes along the path. Back in the work environment the media files and the GPS log can easily be stored with an application (PathCreator) in a hypermedia database, creating a navigable and editable media path (a HyperDocument of the visit). Visitors can load a background picture as a map on which the path is visualised. Printouts with barcodes provide a physical support to navigate the visit. Using multiple projections and physical interfaces, the visitors can re-experience the visit, linking the media material to other physical artefacts (posters, models, objects etc).

Figure 2. The mobile application eDiary on a Pocket PC equipped with the SoapBox as digital compass and with a GPS receiver.

We have developed several components to support the scenario above. The eDiary is a mobile application to run on a PocketPC, which creates a HyperDocument containing a log of time and GPS traces, using a CompactFlash receiver card (Figure 2). The HyperDocument contains HyperNodes for each position; the user can set the recording interval of HyperNodes (e.g. 5 seconds, one minute) short or long, resulting in respectively more or less detailed recorded paths and less or more recorded media in each HyperNode. The GPS coordinates and time log are saved in the meta-information of the HyperNode (Figure 4). The interface of the eDiary displays the current GPS coordinates also on a calibrated map if available, provides a file menu to save the HyperDocument, and has three buttons: start recording; stop recording, and, with an additional button, users can record the direction of a taken photograph by using VTT's SoapBox as an electronic compass (Figure 2). SoapBox (Sensing, Operating, and Activating Peripheral Box) is a light, matchbox-sized device with a processor, a versatile set of sensors, and wireless and wired data communications [16]. With the iPAQ users can also record sounds or write text notes. These files and the created HyperDocument of the visit, containing the time log, the GPS trace, and the compass data, can be saved on an SD (secure digital) memory card. When returning from a visit the user connects the digital camera and the SD memory card to a PC running the

PathCreator. This application is used to combine and synchronise the GPS trace and compass data, which are in the HyperDocument, with the multimedia files: pictures from the camera, sound files, and text files. As each media file has a time of creation, the PathCreator assigns each file to the HyperNode with the nearest time. The HyperNode is a container for multimedia objects, and has position information as we can see from its metainformation in Figure 4, and from its visualisation on a map visualised as a red circle in Figure 3.
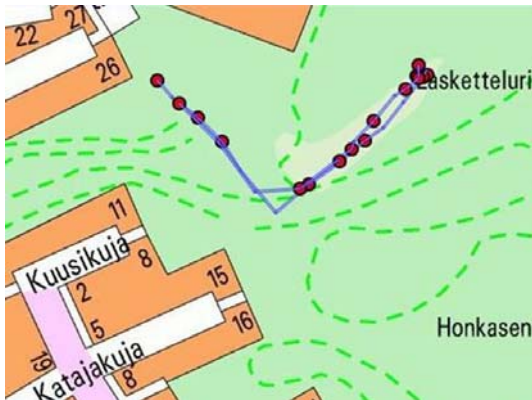

Figure 3. HyperNodes visualized on a map.

Figure 4. Printouts with barcodes can be used to play single pictures or whole *HyperNode*s.

With the iPAQ users can also record sounds or write text notes. These files and the created HyperDocument of the visit, containing the time log, the GPS trace, and the compass data, can be saved on an SD (secure digital) Also the compass data is assigned to the right media object based on the time. For these reasons the clock of digital recording devices needs to be synchronised with the eDiary. The PathCreator also provides loading and GPS calibration of maps, on which the visit can be visualized (Figure 2). After the visit has been saved into the database as a HyperDocument, users can print out HyperNodes with a web browser (Figure 4). Each HyperNode contains a 2D map with a blue trace of the path with a red circle on it positioning the HyperNode. The printouts contain information and thumbnails for all the related objects, and barcodes to play either the whole HyperNode or single objects. The objects can be played on several *MediaPlayer*s on different projections to create an immersive environment.

Configuring the Interfaces and Re-Travelling Visits

For the navigation users have a configurable environment with multiple projections and interchangeable physical interfaces. When a barcode from the print outs is scanned, the corresponding HyperNode or multimedia object is played on the current MediaPlayer. Users can configure specific physical interactions to change the current MediaPlayer or to play the next multimedia object in a HyperNode. The environment is built on an infrastructure with a HyperMedia database where input and output components can register and communicate through XML messaging [14].

Figure 5. the SoapBox has two buttons and communicates through radio signals the data of sensors.

Several physical interfaces can be used:

- Gestures: users can train (configure) their own personal hand gestures by using the gesture recognition system based on VTT Electronics' SoapBox (Figure 5).

- Tilting: users can tilt the SoapBox resulting in four different control commands, front-right, front-left, back-right, back-left.

- Pointing: users can point in eight different direction with a electronic compass as input, north, north-east, east, south-east, south, south-west, west, north-west (e.g. to choose projections or retrieve pictures taken in a direction in a HyperNode).

- Infrared remote control: buttons on a remote control can be associated to events in the environment.

- Barcodes: all commands are available on barcode posters, barcode also turn print outs in physical handles for configurations and digital media.

Alternatively any other physical input can be configured, for example in the Atelier project we used Radio Frequency Identification tags and touch sensors.

Figure 6: Storing and loading configurations for navigation

The hand gestures are detected by the acceleration sensors built into a SoapBox. The measured signal values are wirelessly transmitted from the hand-held SoapBox to a receiver that is connected to a Windows PC with a serial connection. All signal processing and pattern recognition is performed in the PC. The system is capable of detecting discrete 3-dimensional hand gestures and tilting of the device. The rotation (compass bearing) of the device is detected using electronic compass, which is also included in the sensor device [17]. Recognition results can be mapped to different control commands and transmitted using the ATELIER infrastructure. A configuration tool makes it possible to save configurations and load them "on the fly" by scanning a specific barcode (Figure 6). The motivation is to be able, with a barcode, to scan and re-configure the environment according to the preferences of the current user. This was extremely useful in test trials to be able to switch in seconds between different interaction styles. It enabled easy and fast system reconfiguration, for example, two different barcodes could be used for switching from

gesture control mode to pointing mode. A GUI tool provides a way to create, edit and store configuration tables. While most of the interactions in the Atelier environment do not need GUIs, in this case they were needed to make the management of configuration tables easier. The example of a configuration table in Figure 7 shows a configuration where pointing in different directions in the room results in changing the current MediaPlayer (or projection screen) on which the next multimedia object will be displayed (for example by scanning the barcode next to its thumbnail).
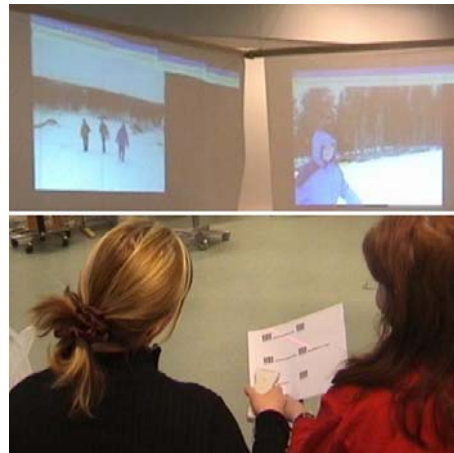

Figure 5. Navigating visits only with the barcode scanner

The creation of the configuration is a mixed interaction between physical interfaces and mouse clicks. Users do not need to write any text, as they first point to a direction with the SoapBox (e.g. south-east). This event results in an entry in the history list (see tab in Figure 6), this entry can be simply clicked and selected. In the same way users can scan barcodes next to projection screens to select specific MediaPlayers (or screens). The corresponding MediaPlayer-screen is then shown on the history list of the Output (Figure 7) and can be selected. Any physical input
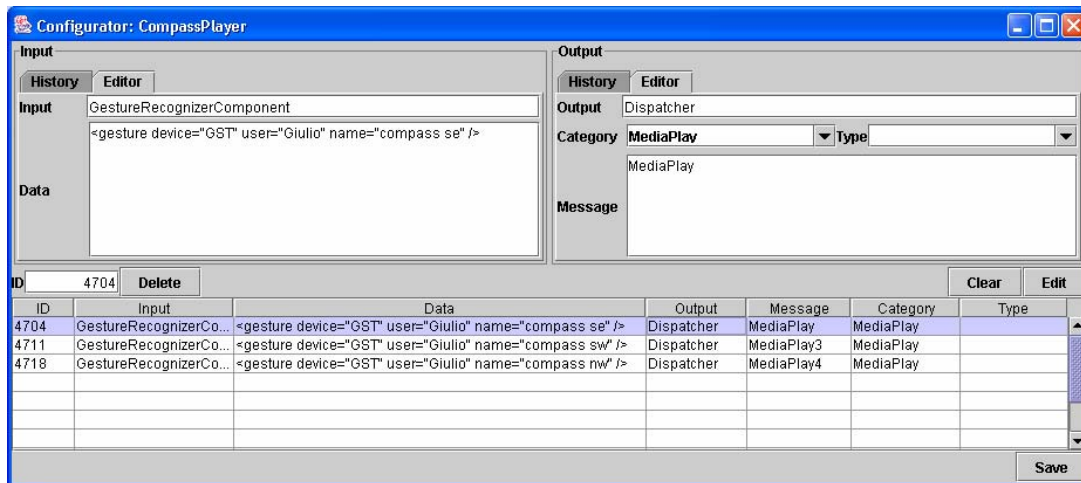

Figure 7. A GUI to edit configurations, events like gestures or scanned barcode or output components, appear in a history list to make configuration a matter of few mouse clicks.

can be therefore configured to trigger specific commands, for example, choose between large projection screens while playing the multimedia recordings from the visits, or moving to next or previous objects in a HyperNode. Pictures are also associated to a direction that can be sensed with the SoapBox as a digital compass during visits. After having scanned the barcode of a HyperNode that contains several pictures, users can also navigate pictures by pointing the SoapBox to different directions and play media that was taken in that direction. This was particularly interesting when during visits a series of pictures form a 360 degrees panorama.

Evaluation in Trials

This analysis is based on a variety of trials that brought us, in the last session, to evaluate the environment as described in the previous chapters. The iterative design resulted in leaving the 3D-map navigation of the path on one of the two projections, adopting a more flexible set up, where the user decides moment-by-moment what to display on each projection screen. Before, to view multimedia recordings users had to move chronologically from one node to another. Now the printouts make the content tangibly available with no fixed scheme of navigation. A 2D map can still be visualised as each HyperNode and relative print out, containing a map showing the position of the HyperNode on the path. In this way the map of each HyperNode is treated as any other multimedia recording. In the last sessions two participants produced three multimedia paths of visits. The objective of this one-day visit was for one participant to show around her town and vicinities to her friend (the other participant). First they visited a skiing facility, producing the first path walking around. Reaching a height by a ski lift, they took panoramic pictures, forming a 360-degree view. Then they visited the town's spa and shopping centre, which were next to each other. Finally the third path was created by car driving to location also outside the town. The two participants were accompanied in their trip by two of our researchers. The participants were equipped with a digital camera while the researchers operated the mobile application eDiary. The participants requested to the researchers to record sounds or spoken notes using the Pocket PC in specific situations. The navigation took place in the Atelier environment arranging the space as depicted in Figure 8. Two back projection

screens were used to display pictures each with a different MediaPlayer. Other two MediaPlayers were used as sound outputs. We organised the trial using three different configurations. The two participants navigated the three visits in the environment, using these three different configurations of physical interfaces, and pictures and sounds triggered discussions and memories of the trip. The participants started to navigate the visits (HyperDocuments) of the previous day, using the infrared remote control and the barcode scanner. These were used to change the active projection screen and move to next or previous objects (Figure 5).
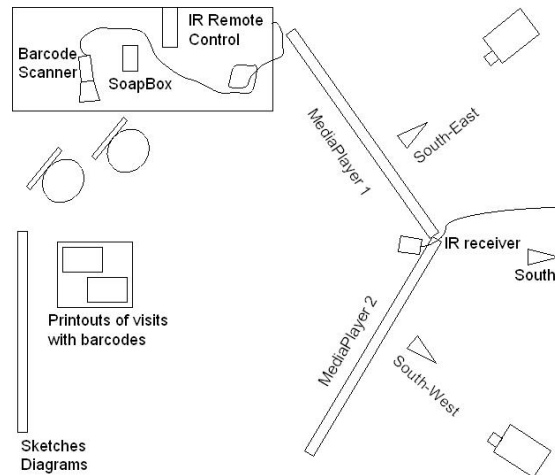

Figure 8. Physical setup of interfaces for the trial.

Then, we loaded a new configuration that enabled to choose the current MediaPlayer using the SoapBox as a compass pointing at different directions (South-East MediaPlayer1, South-West MediaPlayer2, North-East MediaPlayer3 for sound, North-West MediaPlayer4 for sound). Finally, each participant configured own gestures with the SoapBox gesture training system. The gestures for left screen, right screen, and next object, for example, were different from one participant to the other reflecting different personalities and individual ways of gesturing. This third configuration was particularly interesting. Both participants were choosing a MediaPlayer with a gesture, moving the SoapBox from left to right or right to left to respectively choose as current MediaPlayer the South-East and South-West projection (Figure 9). The gesture for next was a
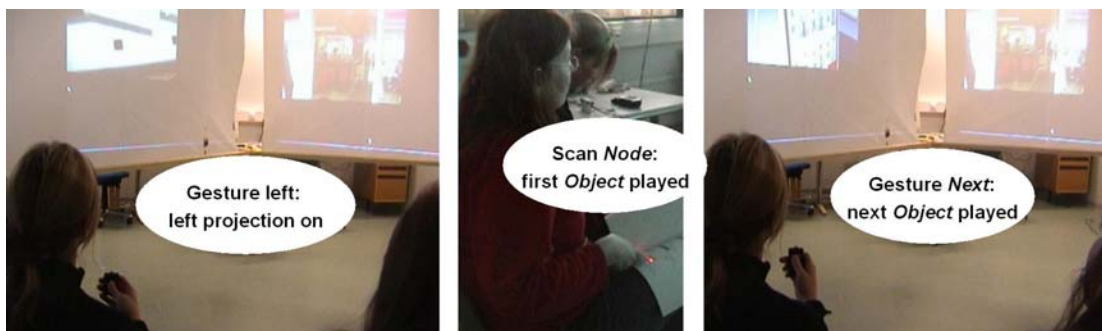

Figure 9. Choosing the left projection and moving to the next objects using gestures.

movement of the Soapbox from the chest towards the front, as using a fishing pole. This resulted in a "throwing" gesture for next, always towards the current screen that had been selected. Although gestures have no orientation (no absolute direction, also relative accelerations) and can be performed in any direction (Figure 9), this feature was particularly useful for allowing the other participants to follow what was happening. Several HyperNodes contained a group of pictures forming a panorama view. Participants could navigate these HyperNodes by first scanning the barcode of the HyperNode and then using the SoapBox as an electronic compass to retrieve pictures that were taken in a specific direction (Figure 10). The availability of several sound outputs (MediaPlayer) resulted in users playing two sounds simultaneously, for examples voices recorded walking outside were mixed together with sounds recorded in a swimming pool resulting in an interesting mixing of "context". The interviews revealed interesting insights on the opportunity to use gesture based interaction in such an environment. Gesture based interaction was experienced as very positive, especially compared to the other modes: "With the gesturing I can see what the other is doing… when she was scanning the barcodes I had to go and check what she was doing … The barcode was boring, and we made mistakes like scanning the same one twice, the gesturing was more active, better." And again one participant explained how she could understand what was going on: "Because this movement (gesture to the left) is so big that you do not have to watch it, you just see it (when the person next to you performs it)." Most of the problems that participants lamented were about the poor feedback of the system: "the feedback was poor to know if the system understood the gestures and if the current player has been changed". This was a problem as: "Sometime I choose a screen then we started talking with Marika and then when I wanted to use the system again I did not remember the current player." Commenting on the print outs as physical handles one participant explained: "It is easy to find what I am looking for, for example when I take a lot of digital pictures now it is hard to find the right picture when I need it because the name is only a number." The use of the SoapBox as an electronic compass was also received very positively, especially in navigating group of pictures

forming a panorama. One participant commented on her navigation of the pictures taken on the height by the ski lift arrival. She observed how the acts of pointing in different directions, discovering again the different pictures, gave her a sense of being there even feeling "the cold weather". The participants also noted how the use of gesture based interaction and the barcode scanner to select the content was possible in collaboration with another person. The SoapBox in our set-up utilises both its buttons: left button activated gesture recognition, the right button activated the use of tilting, and finally pressing both buttons simultaneously enabled the use of the electronic compass. In the discussion with the participants a device was envisioned that would merge all functionality, e.g. a barcode scanner that support gesturing tilting and pointing.

DISCUSSION AND CONCLUSIONS

This paper presented a variety of interfaces and applications to navigate, in an immersive environment, multimedia recordings from visits. The exploration focused explicitly on three opportunities of supporting the viewing of digital pictures in ubiquitous environments: the use of multiple projections and sounds, location information to organise media and re-experience aspects of visits, physical interfaces to render media more tangible and therefore more readily available. We described a distributed system, a mobile application and configuration tools, which embody such a vision. This gave us the possibility to investigate human-computer interaction issues in such an environment. In particular, we have investigated the opportunities and advantages of gesture based interaction. This interaction mode is very promising in such an environment. The results showed that participants configured gestures in different ways, becoming expressive of their own way of gesturing. More importantly, gesture based interaction was particularly helpful to render the interaction intelligible to co-participants. Participants in interview stressed the fact that with gestures they could better follow what was happening and also provide better accountability when the "system" was not responding properly. Finally, it embodied interaction, increasing the active and "physical" participation of users. This, according to participants, was an important feature to re-experience the visit for example
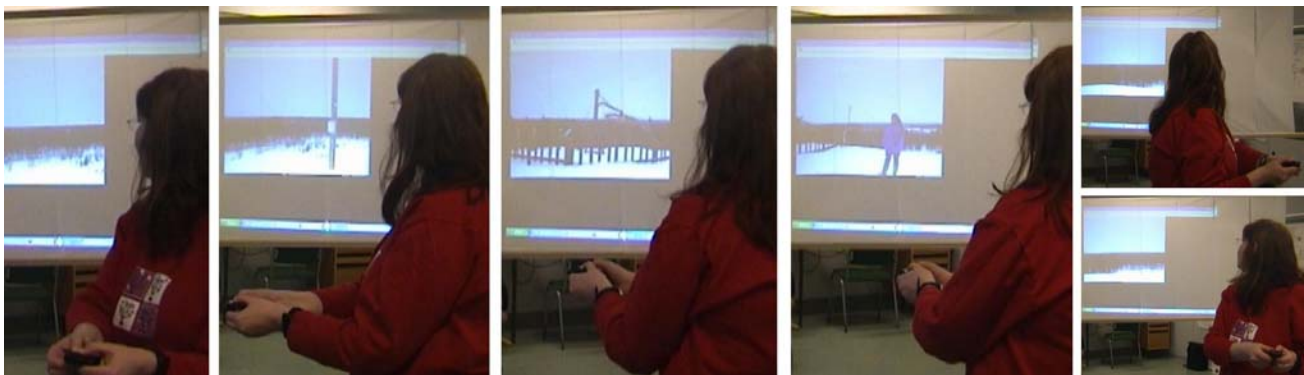

Figure 10. Navigating groups of pictures forming a panorama view using the SoapBox as an e-compass.

by re-discovering pictures pointing in various directions.

REFERENCES

1. Benford, S. Schnädelbach, H. Koleva, B. Gaver, B. Schmidt, A. Boucher, A. Steed, A. R. Anastasi, C. Greenhalgh, T. Rodden, and H. Gellerson. Sensible, sensable and desirable: a framework for designing physical interfaces. Technical Report EQUATOR-03-003, School of Computer Science & IT, Nottingham University, 2003.

2. Davies, N., Cheverst, K., Mitchell, K., Efrat, K., (2001) Using and Determining Location in a Con-text-Sensitive Tour Guide. IEEE Computer 34(8): 35-41.

3. Camurri, A., Volpe G., Gesture-Based Communication in Human-Computer Interaction, 5th International Gesture Workshop, GW 2003, Genova, Italy, April 15-17, 2003, Lecture Notes in Computer Science, 2195, Springer 2004.

4. Camurri, A., Mazzarino, B., Ricchetti, M., Timmers, R., Volpe, G., Multimodal Analysis of Expressive Gesture in Music and Dance Performances. In: Camurri and Volpe (2004).

5. Bellotti ,Victoria and Maribeth Back and W. Keith Edwards and Rebecca E. Grinter and Austin Henderson and Cristina Lopes, Making sense of sensing systems: five questions for designers and researchers, Proceedings of the SIGCHI conference on Human factors in computing systems, 2002, ACM Press.

6. Swindells, C., Inkpen, K.M., Dill, J.C. and Tory, M., (2002) That One There! Pointing to Establish Device Identity. In Proceedings of the 15th annual ACM symposium on User interface software and technology, Symposium on User Interface Software and Technology, 151 – 160.

7. Ailisto H, Plomp J, Pohjanheimo L, Strömmer E (2003). A physical selection paradigm for ubiquitous computing. 1st European Symposium on Ambient Intelligence (EUSAI 2003). Ambient Intelligence. Lecture Notes in Computer Science Vol. 2875. Aarts, Emile et al. (Eds.). Springer-Verlag. Berlin, pp 372 - 383.

8. Sparacino, G. Davenport, and A. Pentland (2000) Media in performance:Interactive spaces for dance, theater, circus, and museum exhibits. In: IBM Systems Journal Vol. 39, Nos. 3 & 4, 2000, p. 479 –510.

9. McGee, Kevin. A Touch of the Future: contact-expressive devices. In: IEEE MultiMedia, 11 (1). January-March 2004.

10. Luff, P., Heath, C. Kuzuoka, H., Hindmarsh, J., Yamazaki, K. and Oyama, S. (2003) 'Fractured ecologies: creating environments for collaboration', Human Computer Interaction 2003, Vol 18 pp 51-84.

11. Balabanovic M, Chu, L.L. & Wolff G. (2000) Storytelling with digital photographs. Proceedings of CHI 2000, 564-571. New York; ACM SIG-CHI.

12. Büscher, Monika, Gunnar Kramp and Peter Gall Krogh, In formation: Support for flexibility, mobility, collaboration and coherence, Personal and Ubiquitous Computing, Vol 7, numbers 3-4, 2003, Springer.

13. Mäkelä, A., Giller, V., Tscheligi, M. and Sefelin, R., 2000. Joking, storytelling, artsharing, expressing affection: A field trial of how children and their social network communicate with digital images in leisure time. Proceedings of CHI'2000. ACM Press, pages 548-555.

14. Iacucci, G., Juustila, A., Kuutti, K., Pehkonen, P., Ylisaukko-oja, A., (2003) Connecting Remote Visits and Design Environment: User Needs and Prototypes for Architecture Design. In: the Proceeding of Mobile HCI 03, 8-11 September 2003, Udine, Italy,Lecture Notes in Computer Science, Springer Verlag. Pp. 45-60.

15. Iacucci, G., Kela, J., Pehkonen, P., (2004) Computational Support to Record and Re-experience Visits, Personal and Ubiquitous Computing Journal, Volume 8, Number 2, Springer Verlag, London. May 2004, Pp. 100–109.

16. Tuulari, Esa; Ylisaukko-oja, Arto, (2002) SoapBox: A Platform for Ubiquitous Computing Research and Applications. Lecture Notes in Computer Sc. 2414: Pervasive Computing. Zürich, CH, August 26-28, 2002. Mattern, F. Naghshineh,M. (eds.). Springer, pp. 125–138.

17. Kallio, S., Kela, J., Mäntyjärvi, J. (2003) Online Gesture Recognition System for Mobile Interaction. IEEE International Conference on Systems, Man & Cybernetics, Volume 3, Oct 5-8, Washington D.C. USA pp. 2070-2076.