Conference contribution :

Robinson, S., Pearson, J. & Jones, M. (in press). *PosterPointing: Making Paper Displays Interactive Using Mobile Devices.* AfriCHI 2016,

PosterPointing: Making Paper Displays Interactive Using Mobile Devices

Simon Robinson, Jennifer Pearson, Matt Jones

[1] FIT Lab, Swansea University, SA2 8PP, UK
{ s.n.w.robinson, j.pearson, matt.jones } @swansea.ac.uk

ABSTRACT

In this paper we explore the potential for creating large pseudo-displays using posters instead of touchscreens. Interactive public displays are common in city centres, but may not be accessible to emergent users. However, the prevalence of increasingly smart, sensor-rich mobile devices offers an alternative. The PosterPointing concept uses physical gestures with mobile devices around printed media in order to explore a rich set of interactions. We created a prototype to investigate the feasibility of one method of enabling this, and suggest other potential approaches. We performed a study to understand how emergent users might gesture using a small device to interact with two larger surface scenarios. Our results show a range of gestures and combinations of phone and paper surface, and offer suggestions for how the design might be used in future.

Author Keywords

Mobile devices; gestural interaction; emergent users.

ACM Classification Keywords

H.5.2 User Interfaces: Input devices and strategies.

INTRODUCTION

Large public displays are used worldwide for a variety of applications, such as information broadcast or on-demand information access. Many of these displays are touchscreens, allowing casual browsing or location-based queries by anyone who approaches the screen. However, these displays are most commonly situated in affluent urban areas, such as tourist attractions, shopping malls or central business districts. As a result, access to—and the potential benefits of—these displays are unavailable to many prospective users.

In this work, we explore whether it might be possible to support similar large display interactions without the need for a display at all. We approached this by considering the technology access and experience levels of the kinds of users who do not currently have the opportunity to use these interactions. It is widely accepted that mobile access is higher than ever before, and that most people worldwide now own or have access to a phone. Increasingly, these phones are smartphones. As Gitau et al. [5] caution us, however, simply having access to a mobile is not the end of the story, and striving towards a future that is mobile-only will leave many behind. Issues around interfaces and interaction designs are especially challenging for emergent users [4].

In situations where GUI-based interactions might not always be appropriate, then, we wondered whether it would be possible to use phones in conjunction with (non-interactive) posters to create a "pseudo-display". Consider the following scenario, which illustrates the concept:

> *Patrick wants to re-listen to a radio programme he heard recently. There's a poster for the radio station on a nearby wall, so he walks over and takes a look, opening the PosterPointing app on his phone. Using his phone as a pointer, he touches different areas, exploring the available content. The phone is able to recognise when he moves the phone to touch each area, turning the poster into an interactive display that he can use to search and filter episodes. After browsing for a while, Patrick finds the presenter he likes, and touches to listen again . . .*

In this paper we present the design concept of PosterPointing. Similar phone+poster interactions have previously been investigated in African contexts – see, for example Smith and Marsden's photo-driven media downloads in taxis [17]. Our primary aim in this work, however was to investigate potential use-cases and interaction approaches for pointing and gesturing *directly* with non-interactive posters. We believe that this is an important step before investing significant research and development time into the idea, so chose to focus on capturing interaction styles, rather than technical developments.

In the rest of this paper, after a summary of related literature, we discuss potential approaches for implementing the concept, including an early technology prototype we have developed. Following this, we present the results of a study conducted to explore the possibility of using phones as pointers with large pseudo-displays. We conclude with a discussion of the potential benefits and challenges of using PosterPointing-like designs in the wild.

BACKGROUND

In this work we focus on two potential approaches to providing phone-as-a-pointer interaction with posters: image-based and sensor-based techniques.

Image recognition and augmented reality

Augmented reality (AR) and image recognition techniques allow digital elements to be layered visually on top of real-world objects, either in real-time or after capturing a photo. Previous work has investigated how these approaches could be used for large display interaction. For example, Ballagas et al. [2] used a grid of visual markers on a large screen in conjunction with a phone camera to allow selection of objects using the phone. Jiang et al. [10] used a similar approach to allow a mobile to move a cursor on a projected display.

In this work, however, we do not aim for interaction with a display; rather, we propose using the mobile as a pointer to interact directly with posters and other printed material. As a result, we cannot use the sorts of synchronisation methods seen in previous work (cf. [2]) Previous work more similar to our concept includes systems such as Anoto [1], which uses specially marked paper and a camera placed adjacent to a pen's nib to recognise document areas and support gestures while writing. Vincent et al. [18] demonstrated AR for tablet-based interaction with a printed wall-mounted map, but focused on optimising map interactions on the touchscreen of the tablet, rather than using the tablet itself as a pointer. Similarly, Henze et al. [8] support poster-based interaction, presenting a useful technique for object recognition, but focusing on interaction on the phone's screen, rather than using its physical form as a pointer. We see this work as being more closely positioned to, for example, Maunder et al.'s SnapAndGrab [14], which used photos of media objects on a large display as surrogates for requesting Bluetooth transfer of related media, or that of Robinson et al. [15], who used a pico-projected surface for pair-based sketching and storytelling.

Tag, token and sensor-based techniques

Previous work has extensively explored using physical tokens to support interaction with and around devices. For example, Rukzio et al.'s early work [16] proposed using NFC or RFID tags to enhance interaction with posters. Hardy et al. [7] built upon this concept to demonstrate how grids of these tags could be used to support the types of pointing interaction we imagine. While this technique has been shown to be effective, our aim is to remove the need for custom-made—and potentially prohibitively expensive—tag grid approaches. Instead, we focus on simplicity and potential widespread utility at the expense of some degree of pointing precision.

As we discuss below, one approach to this could be to use the phone's sensors to detect metal or magnetic objects in the vicinity of the poster the user is interacting with. Previous work has demonstrated how smartphones' onboard sensors can be used to orient objects around the device in this manner. For example, Ketabdar et al. [11] showed how a magnet held above a smartphone could be used to gesture to enter digits. Liang et al.'s approach [12] used a grid of magnets attached to a phone to provide accurate stylus input (supporting tilt, pressure and other features). Other related work includes that of Bianchi and Oakley [3], who demonstrated on-device and adjacent-to-device tangible objects using magnetic fields.

The key difference between our concept and the majority of previous work is that prior research has focused interaction

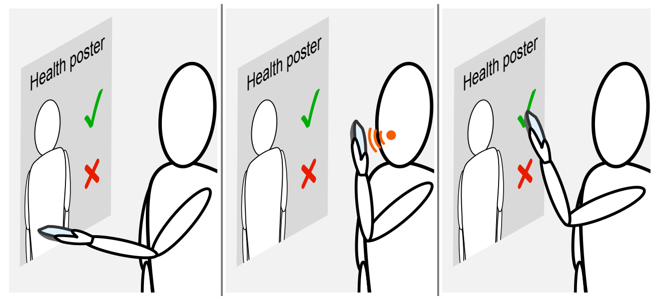

Figure 1. PosterPointing concept. From left to right: Senzo points to symptoms on the new self-help health information poster, then listens to the advice given. Wanting an expert's opinion, he records a short description of his symptoms by speaking into his phone, then points to the poster again to send details to a doctor for diagnosis.

around or on the device detecting movement. That is, whether using AR [6], NFC [13], multi-tag gestures [9] or other tracking approaches, related work has used the poster interaction as a way to initiate or support tasks that are then completed on the touchscreen of a smartphone or other device. In our work we imagine moving the detecting device (the smartphone), and using it as a pointer to other objects, offering a coarser level of precision but a potentially more natural interaction.

CONCEPT AND PROTOTYPE

As discussed in the previous section, there are various technological approaches that could potentially support the type of interactions we imagine. Figure 1 illustrates how we see the system being used in a self-help medical context, for example. Due to the nature of the design and the sensing techniques available, the degree of pointing precision that we are aiming for is lower than previous work. However, rather than supporting very precise interactions using pointing alone, we imagine using the system in conjunction with other multimodal feedback (activating relevant audio content over a low-bandwidth connection, for instance). Furthermore, movements over the surface could provide "drag and drop" and other functionality, and combinations of mobile screen interactions and surface movements may provide even richer interactions, such as pinch to zoom.

We have developed a prototype that supports some of these interactions. Currently, our design uses smartphone sensors in conjunction with a neodynium magnet positioned below a QR code in the centre of an A4-sized poster. After scanning the QR code (to uniquely identify the poster), our prototype prompts the user to perform a calibration step, by pointing to all four corners of the poster with the corner of their phone (similar to [15]). This calibration information is saved for future interaction with this particular poster. Subsequently, the prototype is able to detect when the user touches the same corner of the phone elsewhere on the poster.

Our current design can support gross selections, detecting up to approximately ten unique areas on the poster. While this is potentially sufficient for some of the concept scenarios we imagine, its level of accuracy is not yet high enough for the richer interactions we have designed. Consequently, we are now turning our attention towards image recognition-based methods, as described in the background section. We imagine

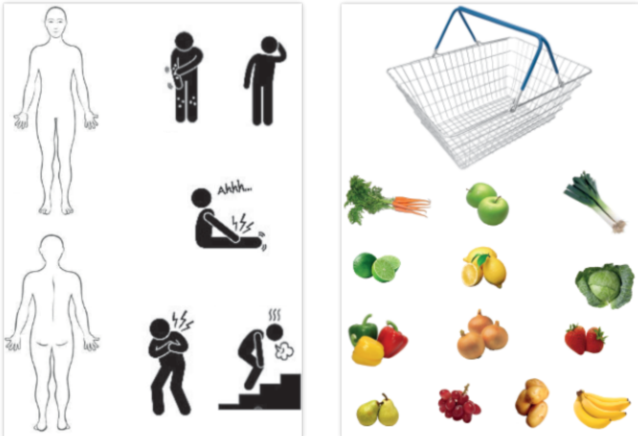

Figure 2. Interaction surfaces used in the study. Left: the medical poster, showing the body (front and back) to the left of the image, and example symptoms to the right of the image. Right: the shopping scenario, with a collection of items available to buy.

a refined prototype working in much the same way as our current design, but using the back of the phone rather than the corner as the pointer (e.g., keeping the poster in view of the camera); and, without the need for any initial user calibration.

Before carrying out this work, however, we felt it appropriate to explore the types of interactions and gestures that might be used by potential users, allowing us to assess, at a basic level, the degree of accuracy that the system would need to support.

EXPLORATORY STUDY

We performed an IRB-approved lab study in Mumbai, India in order to explore the potential viability of the approach, and its resonance with emergent users. Eight people were recruited to take part in a 20 min prototyping session. Participants were aged between 18–27, and all were male, working as housekeepers or casual labourers. All participants owned phones: one participant owned a smartphone, another owned a featurephone, and the remainder owned basic phones. Seven of the participants were classified at the Rote Learner stage of Devanuj and Joshi's User-Usage Model (cf. [4]), with the sole smartphone user classified as Fluent.

The focus of this study was twofold. Firstly, we wanted to establish the viability of using the phone as a gestural pointer for users who are not familiar with advanced mobile devices and services. Our second goal was to explore the types of gestures and movements participants might choose to use with a small device in conjunction with a larger touch surface, allowing us to assess whether a future version of our prototype could support the sorts of interactions participants imagine using.

Procedure

At the start of the study participants were welcomed, given information about the purpose of the study and its tasks, and guided through an ethical consent procedure to ensure they were fully aware of their rights during the study, including drawing particular attention to our request to record video of the gestures they used. After participants consented, we collected basic demographic and technology experience information in addition to recording their level of education.

Following this, participants were asked to perform gestures to interact with two separate interaction scenarios printed on wall-mounted posters (see Fig. 2). The first of these was a self-help medical diagram (Fig. 2, left), with participants asked to imagine that the poster interaction would be used as part of an early informational or self-diagnosis procedure before visiting a doctor. The second scenario involved grocery shopping (Fig. 2, right), with participants thinking of ways to use gestures to indicate items that they wished to buy.

To demonstrate their gesture ideas, participants were given a non-functional (i.e., powered off) smartphone, and told that they should use the phone to interact with the poster. The phone was powered off in order to focus participants' attention on the device as a passive, location-tracking pointer rather than on its technological capabilities (which have been amply demonstrated in previous work in this area). Participants were told that they could touch anywhere on the phone or on the paper (or both in tandem) and that the phone knew where they were touching and how they were holding it.

The study tasks were semi-structured, beginning with a request from the researcher (e.g., *"Imagine you had a rash on your stomach. How would you show your phone what is wrong with you?"*). The participant then used the phone in conjunction with the poster to indicate—through touching, tapping or any other method of their choice—how they would indicate this condition. After the participant had demonstrated their gestures for this action, they were asked to show other variations and responses (e.g., *"How would you show that you have more than one symptom?"*), to explore the range of gestures that they might use to interact. The same format was used for the shopping scenario, with participants asked to buy a number of items, and then return one or more they no-longer wanted.

These tasks were repeated a second time (e.g., all participants interacted with a second set of both posters), with participants asked this time to briefly describe why they chose particular gestures, or whether they had any alternative gestures in mind. Finally, participants were thanked for their time, and given ₹100 as compensation for their participation.

RESULTS

We analysed the videos of each participant's interactions, paying particular attention to the areas of the phone that were used, gesture styles (e.g., tapping, pointing, drawing etc.), and any compound interactions that participants employed (such as using both pointing and another modality to interact). These observations were then clustered into distinct interaction forms, which we present below.

Participants demonstrated a wide range of gestures and interaction styles when interacting with the two scenarios. The most common approach to the tasks was for participants to touch the corner of the phone to a point on the poster to select items or positions. Six of the eight participants took this approach (see Fig. 3, (a–b)). Of the remaining participants, one used the top edge of the phone to touch in a similar manner to those who had used the corner, while the last participant used the whole phone to gradually zoom into items to select them, as if using the camera to recognise the images, as shown in Fig. 3, (c).

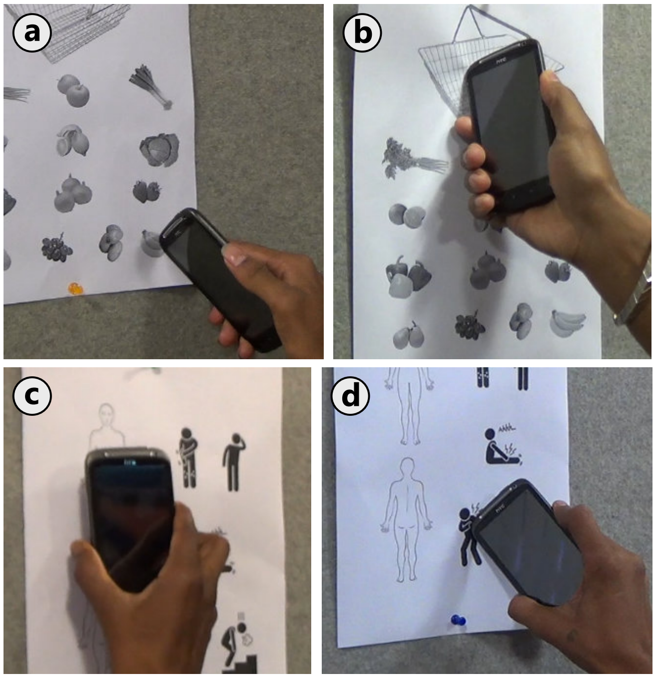

Figure 3. Gesture styles. The most common method was to use the corner of the phone, either tapping (a) or drawing lines (b). Other styles were also observed, including zooming in to the image as if using the camera to select (c). Participants sometimes customised gestures to add context – for example, by pressing harder to indicate feeling more pain (d).

Gesture styles

Participants tended to use one of two broad interaction styles observed: tapping the phone to positions in turn, or drawing a line between two or more positions. Half of the participants used each approach. Participants customised their gestures depending on the scenario that was in use, however, and several participants used both styles.

Self-help medical scenario

In the medical scenario, participants tended to touch part of the body and then touch the symptom image (in either order), indicating both the location and type of ailment that they were reporting. There were interesting behaviours around customising this response, however. For example, one participant wanted to indicate particularly severe pain, so they rubbed the corner of the phone on the symptom image, pressing especially hard to indicate the severity of the condition (see Fig. 3, (d)). Another participant combined modalities, choosing to ignore the symptoms listed, and instead touching a body position and saying the word for the symptom. Finally, one participant chose not to touch the symptom images, but instead drew a tick or cross gesture in the air over each one, and then touched the body position to indicate where the condition was affecting them.

Shopping scenario

The shopping scenario highlighted several different behaviours in terms of item selection. The most common interaction was for participants to touch items and then touch the basket, using this gesture to indicate which items they wanted to buy. For quantities, participants repeated the whole gesture (e.g., three times for three items).

One interesting observation came when participants were asked to buy multiple different items, or all of the items at once. In this case, all participants chose to draw a line from item to item, making sure to touch each one, and then either drew a line to or touched the basket. None of the participants used the spatial arrangement of the items to, for example, allow circling or multiple selections – it was seen as important to touch all the items to select them.

When asked to remove items from the basket, six participants reversed the gesture they had used previously, though one participant touched the item and then made a gesture in the air (e.g., to throw the item away), and another touched the item and then swiped the screen of the phone.

Findings

Previous work in this space has largely focused on completing the final step of precise tasks using a touchscreen (e.g., [18]), but our participants focused primarily on gestures with the phone itself. Clearly, our experimental design encouraged participants to use the phone in conjunction with the poster, but we would argue that the range of similar gestures used by participants suggests a space for the types of interaction we are investigating.

Interestingly, some participants chose to demonstrate multimodal gestures, entirely unprompted. This finding offers support for our concept of interacting with, say, an audio service, or potentially another way to refine gesture accuracy.

In general, participants' gestures replicated the standard task flow of physical versions of the tasks (e.g., touching an item before the basket to put the item in; reversing this to remove the item). Participants did not take advantage of the spatial arrangement of the poster to select multiple items, or perform larger-scale gestures, such as circling a group of items. This suggests that future versions of our design will need higher precision than is currently supported. While most participants naturally used the corner of the phone to touch items, one participant chose to hold the phone as if its camera was viewing the item being selected, suggesting that camera-based tracking could be used for a refined version of our prototype.

CONCLUSION

In this work we have explored the potential for creating pseudo-displays using posters rather than touchscreens. We asked emergent users to carry out a range of simple transactions using a phone+poster interaction technique of their choosing. Seven of eight participants did not have access to smartphones, and none had previously interacted with sensor-based interfaces. Despite this, all were able to imagine and demonstrate interacting with a poster in this way, generating a range of different gestures and combinations of the phone and the paper display.

Our early prototype of the PosterPointing system uses inertial sensor-based tracking to provide a relatively low level of accuracy. The next step in this work is to develop a refined prototype, initially using camera-based tracking. We then plan to deploy the design with emergent users in context.

ACKNOWLEDGEMENTS

This work was funded by EPSRC projects EP/M00421X/1 and EP/J000604/2. We would like to thank Anirudha Joshi, Abhishek Shrivastava, Manjiri Joshi, Nagraj Emmadi, Nirav Malsattar and other colleagues for their help in coordinating the studies described in this work.

REFERENCES

1. Anoto AB. 2011. Anoto technology. http://anoto.com.
2. Rafael Ballagas, Michael Rohs and Jennifer G. Sheridan. 2005. Sweep and point and shoot: phonecam-based interactions for large public displays. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '05). ACM, New York, NY, USA, 1200–1203. http://dx.doi.org/10.1145/1056808.1056876.
3. Andrea Bianchi and Ian Oakley. 2013. Designing tangible magnetic appcessories. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction* (TEI '13). ACM, New York, NY, USA, 255–258. http://dx.doi.org/10.1145/2460625.2460667.
4. Devanuj and Anirudha Joshi. 2013. Technology adoption by 'emergent' users: the user-usage model. In *Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction* (APCHI '13). ACM, New York, NY, USA, 28–38. http://dx.doi.org/10.1145/2525194.2525209.
5. Shikoh Gitau, Gary Marsden and Jonathan Donner. 2010. After access: challenges facing mobile-only internet users in the developing world. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '10). ACM, New York, NY, USA, 2603–2606. http://dx.doi.org/10.1145/1753326.1753720.
6. Jens Grubert, Raphaël Grasset and Gerhard Reitmayr. 2012. Exploring the design of hybrid interfaces for augmented posters in public spaces. In *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design* (NordiCHI '12). ACM, New York, NY, USA, 238–246. http://dx.doi.org/10.1145/2399016.2399053.
7. Robert Hardy, Enrico Rukzio, Paul Holleis and Matthias Wagner. 2010. Mobile interaction with static and dynamic NFC-based displays. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services* (MobileHCI '10). ACM, New York, NY, USA, 123–132. http://dx.doi.org/10.1145/1851600.1851623.
8. Niels Henze, Torben Schinke and Susanne Boll. 2009. What is that? object recognition from natural features on a mobile phone. In *Proceedings of the 2009 Workshop on Mobile Interaction with the Real World* (MIRW '09). http://nhenze.net/uploads/What-is-That-Object-Recognition-from-Natural-Features-on-a-Mobile-Phone.pdf.
9. Rosen Ivanov. 2015. A novel approach for accessing web services based on multi-touch interaction with NFC smart posters. In *Proceedings of the 16th International Conference on Computer Systems and Technologies* (CompSysTech '15). ACM, New York, NY, USA, 152–159. http://dx.doi.org/10.1145/2812428.2812433.
10. Hao Jiang, Eyal Ofek, Neema Moraveji and Yuanchun Shi. 2006. Direct pointer: direct manipulation for large-display interaction using handheld cameras. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '06). ACM, New York, NY, USA, 1107–1110. http://dx.doi.org/10.1145/1124772.1124937.
11. Hamed Ketabdar, Mehran Roshandel and Kamer Ali Yüksel. 2010. Magiwrite: towards touchless digit entry using 3D space around mobile devices. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services* (MobileHCI '10). ACM, New York, NY, USA, 443–446. http://dx.doi.org/10.1145/1851600.1851701.
12. Rong-Hao Liang, Kai-Yin Cheng, Chao-Huai Su, Chien-Ting Weng, Bing-Yu Chen and De-Nian Yang. 2012. Gausssense: attachable stylus sensing using magnetic sensor grid. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (UIST '12). ACM, New York, NY, USA, 319–326. http://dx.doi.org/10.1145/2380116.2380157.
13. Antonio Lotito, Giovanni Luca Spoto, Antonella Frisiello, Vito Macchia, Thomas Bolognesi and Francesco Ruà. 2013. Smart2poster. bridging information and locality. In *Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services* (MobileHCI '13). ACM, New York, NY, USA, 582–587. http://dx.doi.org/10.1145/2493190.2494440.
14. Andrew Maunder, Gary Marsden and Richard Harper. 2007. Creating and sharing multi-media packages using large situated public displays and mobile phones. In *Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services* (MobileHCI '07). ACM, New York, NY, USA, 222–225. http://dx.doi.org/10.1145/1377999.1378010.
15. Simon Robinson, Matt Jones, Elina Vartiainen and Gary Marsden. 2012. Picotales: collaborative authoring of animated stories using handheld projectors. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (CSCW '12). ACM, New York, NY, USA, 671–680. http://dx.doi.org/10.1145/2145204.2145306.
16. E. Rukzio, A. Schmidt and H. Hussmann. 2004. Physical posters as gateways to context-aware services for mobile devices. In *Sixth IEEE Workshop on Mobile Computing Systems and Applications, 2004* (WMCSA '04), 10–19. http://dx.doi.org/10.1109/MCSA.2004.20.

17. Graeme Smith and Gary Marsden. 2011. Providing media download services in African taxis. In *Proceedings of the South African Institute of Computer Scientists and Information Technologists Conference on Knowledge, Innovation and Leadership in a Diverse, Multidisciplinary Environment* (SAICSIT '11). ACM, New York, NY, USA, 215–223. http://dx.doi.org/10.1145/2072221.2072246.

18. Thomas Vincent, Laurence Nigay and Takeshi Kurata. 2013. Precise pointing techniques for handheld augmented reality. In *Human-Computer Interaction – INTERACT 2013: 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part I* (INTERACT '13). Springer Berlin Heidelberg, Berlin, Heidelberg, 122–139. http://dx.doi.org/10.1007/978-3-642-40483-2_9.