Move Your Phone: Spatial Input-based Document Zoom & Pan on Mobile Displays Revisited

Martin Spindler
User Interface & Software Engineering Group
University of Magdeburg
Germany
martin.spindler@acm.org

Marcel Martsch
Dep. of Vocational Education & Human Resources Development
University of Magdeburg
Germany
marcel.martsch@ovgu.de

Martin Schuessler
Quality and Usability Lab
Telekom Innovation Laboratories
TU Berlin, Germany
schuesslerm@acm.org

Raimund Dachselt
Interactive Media Lab
Technische Universität Dresden
01062 Dresden, Germany
dachselt@acm.org

CHI 2014, Apr 26 - May 01 2014, Toronto, ON, Canada
ACM 978-1-4503-2474-8/14/04.
http://dx.doi.org/10.1145/2559206.2574777

Please Note: *This is an Extended Abstracts entry for an Interactivity exhibit at CHI 2014. It accompanies a fully refereed article that can be found in the CHI 2014 Main Proceedings.*

Abstract

We present a document navigation technique for mobile displays that relies entirely on principles of spatial manipulation, such as lifting the display up/down to zoom. While the underlying concepts are not new, the goal of this interactivity is to demonstrate the potential of spatial input-based navigation on state-of-the-art mobile displays. For this purpose, we implemented two carefully optimized prototypes using popular consumer hardware (iPhone and iPad). We originally developed these prototypes for a comprehensive user study [4], in which we found overwhelming proof that spatial manipulation can – if designed and implemented properly – outperform conventional multi-touch-based 2D document navigation. These findings could be of interest for future interaction designs of mobile devices. With this interactivity, we want to share our hands-on experiences with the CHI community.

Author Keywords

Spatial Input; Spatially Aware Displays; 2D Document Navigation; Pinch-Drag-Flick; Mobile Displays

ACM Classification Keywords

H.5.2. Information interfaces and presentation: User Interfaces – *input devices and strategies*.

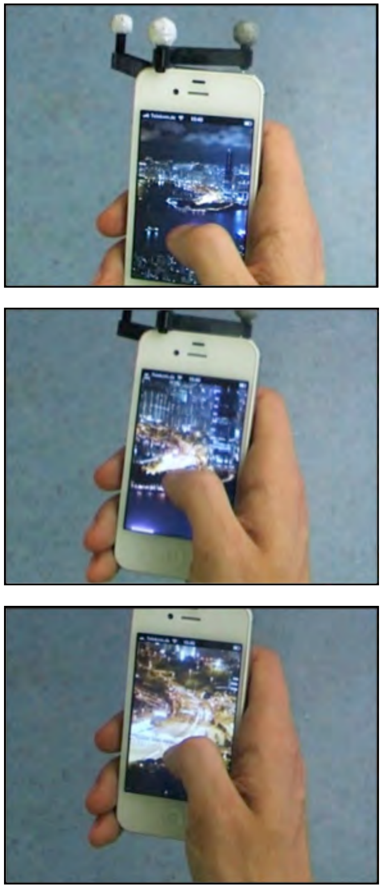

Figure 1 (iPhone): Users can zoom into a picture by moving the display along its local Z-axis. As spatial input is disabled by default, it must be activated explicitly by touching the screen (clutching).

Introduction

The exploration of large 2D information spaces, such as maps, pictures and web documents, is a common task carried out on mobile displays by millions of users every day. Due to the rather small screen size of the devices, this often involves heavy usage of zoom and pan, usually performed using multi-finger gestures. In this context, the Pinch-Drag-Flick paradigm has proven to be one of the most (commercially) successful gesture sets: pinch to zoom, drag and flick to pan. While these gestures are considered to be easy to learn and perform, there are inherent problems with the approach: fingers occlude items on the screen; travel distances per gesture are short; pinch gestures are difficult to execute if one hand is occupied; and elderly or disabled persons may not possess sufficiently fine motor skills to perform gestures accurately. In this interactivity, we demonstrate a radically different navigation approach that is based on an alternative input channel: the spatial position and orientation of mobile displays, e.g., as proposed in [1] and [5]. In contrast to the metaphor of grabbing a document, spatial input-based navigation uses the metaphor of moving a viewport (a display) over a virtual information world. For this purpose, distinct motion patterns are mapped to specific navigation tasks, e.g., horizontal movements may change the viewport center (panning), whereas lifting a display up/down may control the zoom factor. As this requires users to move a display through the physical space surrounding them, the motor space is increased considerably and a different set of motor skills is addressed. We see this difference in motor control as a significant opportunity that may help overcome the problems of conventional touch-based navigation – not as a superior form of interaction, but as a complementary one.

Surprisingly little practical work has been done on implementing (and studying) how these approaches perform against each other on state-of-the-art mobile displays. In fact, the few prior attempts (see [2] and [3]) could not demonstrate the true potential of spatial-based navigation, e.g., in terms of efficiency. This might be attributed to specific design decisions, e.g., no use of clutching, lack of state-of-the-art technology (e.g., cable-bounded devices), and probably a low quality implementation (the latter is speculative, yet based on our own experiences as we spent long hours on making the spatial technique as robust and responsive as conventional Pinch-Drag-Flick).

With our work, we contribute a high quality implementation of the spatial approach using popular consumer hardware (iPhone 4 and iPad 3). These prototypes can be considered as a “byproduct”, as we originally designed and built them for a comprehensive user study [4]. In this study, the 40 participants were on average more than 35% faster with the spatial technique, even though all of them were conversant with Pinch-Drag-Zoom and used the spatial approach for the first time. We believe that this is due to several design decision that we made, e.g., regarding the importance of an easy to use clutch and the role of a high quality prototype, as will be discussed next.

Design Decisions

Mapping the Physical to the Virtual World

One key question is how to properly map the physical to the virtual space. We tried various mappings and decided on a dynamic mapping that uses the current orientation of the display as the new reference plane for future interpretations of motions. This means that zooming is mapped to movements along the normal of

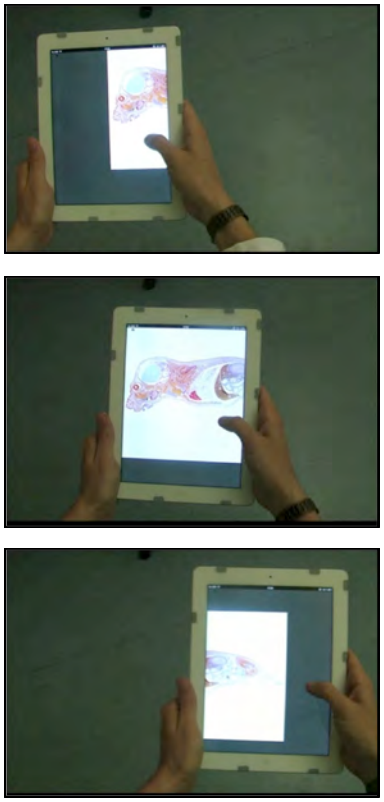

Figure 2 (iPad): Panning is done by moving the display sideways. Pan motions are stopped automatically when document boundaries are about to cross the middle of the screen (the zoom center lies there).

the display (local Z-axis), whereas motions within the display's XY-plane define panning. Our experiences show that a dynamic mapping supports body-centric usage better than spherical mappings [1]. It also has the benefit of working independently of the user's position, thus simplifying the interaction design and spatial tracking.

Clutching and Relative Mode

Most of the time, mobile displays are moved without any intention to interact. Thus, spatial input should be inactive by default, only to be *enabled on purpose* for a brief moment of interaction – by activating a clutch. With a clutch, the nature of spatial navigation can be changed from absolute to relative mode. We believe that this is a very important and necessary step to support mobile usage. In relative mode, the "volumetric" 2D document (represented by a pyramid Space-Scale-Diagram) travels along with the device like a bubble surrounding it. This enables users to put away the device and to resume navigation later on with the last visited position. While we fiercely advocate the use of tactile clutches, we decided on a touch-based variant, primarily due to practical reasons (see [4]). Users can activate the clutch by touching the screen with one or more fingers, e.g., close to the screen bezel in order to prevent occlusion of items in focus. This enables users to quickly access the clutch without spending much mental effort on locating it.

Zoom Center and Direction

The zoom center remains fixed to the middle of the screen and users cannot modify its position. Backed by a pre-test with 5 users, we decided to zoom out, when the device gets closer to the body. This can be switched if the opposite direction is preferred.

Pan Boundaries

Special care must be taken for handling document boundaries to prevent users from navigating into the void. This is usually addressed by stopping pan motions at the document borders. We adjusted this behavior so that users can align even the document corners to the zoom center (the middle of the screen, see Figure 2).

Implementation

For spatial tracking, we use 6 infrared (IR) cameras mounted to a portable traverse. This enables us to precisely determine the spatial position and orientation (6DOF) of iPhone and iPad at 30Hz with an error of less than 1mm within the tracking volume covering a projected area of about 3m × 3m. We attached 6 unobtrusive IR-stickers to the display bezel of the iPad (see Figure 2). Only 3 of them need to be visible at a given time. This allows users to hold the device freely in their hands without accidentally interrupting the tracking. On the iPhone, we use a small lightweight plastic frame with 4 IR-reflective spheres that can be plugged into the iPhone's headphone output (see Figure 1). Spatial raw positions and orientations of iPhone and iPad are continuously streamed over a local Wi-Fi network to the devices in a standardized form (VRPN).

We implemented the prototype in Objective C using the native development tools of iOS 6.0 (XCode 4.5). A major problem was the limited RAM of the iPhone (500 MB) of which we often got assigned less than 100 MB by the system (on which we had little influence on, even though we used a fresh system with no additional apps installed). Once loaded, an uncompressed image (16 Megapixel × 32 color bits = 64 MB) already consumed most of the available memory. This led to frequent crashes due to insufficient memory, which also

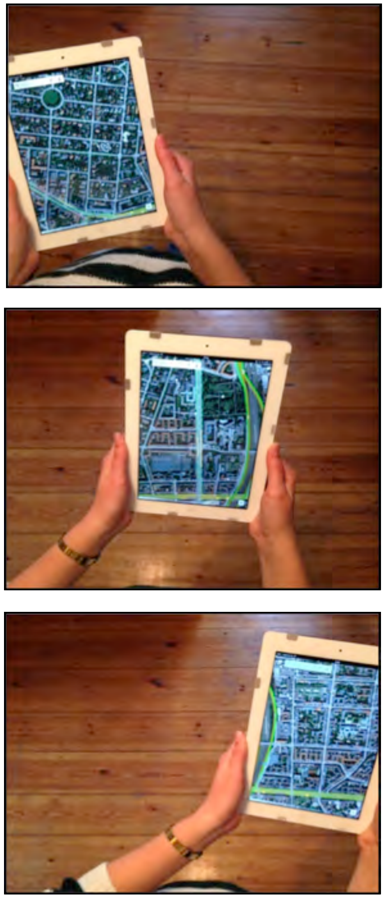

Figure 3 (Map Navigation): Because the map is virtually endless, no special treatment of document boundaries is necessary. This also increases the use of clutching (not demonstrated in the pictures).

occurred on the iPad (1GB RAM), yet less frequently. We encountered massive performance fluctuations for the spatial technique, mostly caused by the large travel distances per gesture. A single lift of the device, e.g., required the algorithm to quickly swap between high-resolution details and low-resolution overviews. For touch, in contrast, there was usually much more time for fetching neighboring content, as the travel speed per gesture was slower. Beyond that, the iOS software is also heavily optimized towards touch input. After testing various data representations, we found a solution that combines several strategies: A zoom pyramid consisting of three layers containing different resolutions of the picture, with the most detailed one being built up of tiles that are loaded on demand.

Example Applications

To enable users to test and experience our thoroughly optimized implementation of spatial input-based navigation on the iPhone and iPad, we designed two example applications that we demo at CHI. Both apps also support conventional Pinch-Drag-Flick navigation, allowing for a direct comparison of the two approaches.

Picture Viewer

In the first demo, users can zoom and pan through several high-resolution images, e.g., a photo of Hong Kong at night (see Figure 1) and a cut through a rat embryo (see Figure 2).

Map Navigator

The second demo demonstrates map navigation (see Figure 3). It is based on the maps service of iOS and showcases the exploration of unbounded (open) information worlds, which is accompanied by more frequent clutching.

Conclusions

We built a high quality prototype demonstrating the potential of spatial input-based document navigation on state-of-the-art mobile displays. Unlike previous attempts, our implementation outperforms conventional touch-based Pinch-Drag-Zoom. This was confirmed by a comprehensive user study that we conducted [4]. Given the additional advantages of a supplemental input channel, we hope that our findings help mobile computing embrace spatial interaction principles much more than before – as a further method of interaction, yet not as a replacement. Because there are also limitations: social protocols may limit its application, users may perform differently when sitting, or they may prefer to put a display on a desk for certain tasks.

References

[1] Fitzmaurice, G.W., Zhai, S., and Chignell, M. Virtual reality for palmtop computers. *ACM Trans. Inf. Syst. 11*, 3 (1993), 197–218.

[2] Hansen, T.R., Eriksson, E., and Lykke-Olesen, A. Mixed interaction space – Expanding the interaction space with mobile devices. In *Proc. HCI 2006*. Springer London (2006), 365–380.

[3] Pahud, M., Hinckley, K., Iqbal, S., Sellen, A., and Buxton, B. Toward compound navigation tasks on mobiles via spatial manipulation. *Proc. Mobile HCI 2013,* ACM (2013), 113–122.

[4] Spindler, M., Schuessler, M., Martsch, M., and Dachselt, R. Pinch-Drag-Flick vs. Spatial Input: Rethinking Zoom & Pan on Mobile Displays. In *Proc. CHI 2014*, ACM (2014).

[5] Yee, K. Peephole Displays: Pen Interaction on Spatially Aware Handheld Computers. In *Proc. CHI 2003*, ACM Press (2003), 1-8.