ACTUI: Using Commodity Mobile Devices to Build Active Tangible User Interfaces

Ming Li
Computer Graphics & Multimedia Group
RWTH Aachen, AhornStr. 55
52074 Aachen, Germany
mingli@cs.rwth-aachen.de

Leif Kobbelt
Computer Graphics & Multimedia Group
RWTH Aachen, AhornStr. 55
52074 Aachen, Germany
kobbelt@cs.rwth-aachen.de

MobileHCI '15 Adjunct, August 24-27, 2015, Copenhagen, Denmark

http://dx.doi.org/10.1145/2786567.2792895

Abstract
We present the prototype design for a novel user interface, which extends the concept of tangible user interfaces from mostly specialized hardware components and studio deployment to commodity mobile devices in daily life. Our prototype enables mobile devices to be components of a tangible interface where each device can serve as both, a touch sensing display and as a tangible item for interaction. The only necessary modification is the attachment of a conductive 2D touch pattern on each device. Compared to existing approaches, our Active Commodity Tangible User Interfaces (ACTUI) can display graphical output directly on their built-in display paving the way to a plethora of innovative applications where the diverse combination of local and global active display area can significantly enhance the flexibility and effectiveness of the interaction. We explore two exemplary application scenarios where we demonstrate the potential of ACTUI.

Author Keywords
Tangible; Mobile Devices; Touch Screen; Touch Pattern; Pose Tracking; Magic Lens; Bench Viewer; UI Design

ACM Classification Keywords
H.5.2 [Information interface and presentation]: User Interfaces - Graphical user interfaces

Figure 1: Example application: Magic lens. The stacked display can be used to show an extra information layer, like a synchronized satellite image over a traffic map.

Figure 2: Example application : Bench viewer visualizes a 3D volume data.

Introduction

In the quest for ever more effective, efficient, and intuitive user interfaces to explore and manipulate visual content, there was a long evolution from real desktops to virtual desktops to touch sensitive displays and finally to Tangible User Interfaces (TUI) where physical props are used as input devices to control a virtual application being displayed on an underlying touch sensitive table top. However existing solutions usually require a complex studio setup and specialized hardware components, which reduce the accessibility and flexibility for users in daily life. In this paper we want to explore the design of TUIs in mobile application scenarios, where we apply active tangible components (mobile devices) that have their own (local) input and output facilities. The opportunities emerging from the combination of local and global display areas considerably extend the flexibility in the layout of interaction techniques since, e.g., each display can show a well-coordinated and synchronized view on a common 2D or even 3D dataset. To avoid the development and manufacturing of specialized hardware, we build a prototype for our Active Commodity Tangible User Interfaces (ACTUI) from commodity mobile devices that are equipped with capacitive multi-touch displays by default (e.g. iPhones and iPads).

The basic idea of our approach is that through its capacitive multi-touch display a mobile device placed on the table can detect and track the identity and pose (location and orientation) of another device stacked on top of it if the stacked device is equipped with a unique conductive 2D touch pattern on its backside. This relative pose information allows the system to properly align the viewports of both devices and thus to correctly synchronize the content shown on both devices. One obvious application for this functionality is the implementation of a magic lens that displays a zoomed-in view or another information layer on the stacked device (see Figure 1). There are many more potential ways to exploit this functionality, like a non-planar setup, bench viewer (see Figure 2), leveraged by the possibility to place one device perpendicularly on the other (and not just flat).

Related Work

Tangible User Interfaces (TUI) are generally known as interacting with virtual digital data via physical controllers. A good example is Siftable [6], a tangible user interface consisting of multiple small smart devices which can display graphical content and detect neighboring devices. Another example is display blocks [10], which builds a cube display using 6 small screens to visualize different perspectives of a virtual object. Such TUIs usually require specific devices that are not broadly available to casual users. The limited computing power and display resolution also restrict its functionality. TUIs have also been developed for table-tops [4, 7, 11, 3]. Weiss et al. presented SLAP widgets [17] for tangible multi-touch tabletops, where they used transparent props as input devices that can change their appearance by using the underlying display in a see-through fashion (e.g. keyboards). These approaches provide big display surface and high computing power for interaction, which makes it fit very well in the context of ubiquitous computing, but they still need special devices that are either heavy, expensive or complicated to deploy.

In order to enable TUI on mobile devices, many research efforts have been invested into the usage of commodity mobile devices and new sensing techniques. Yu et al. [19] proposed two different touch patterns: spatial tag and frequency tag. A tangible object's identity is encoded in either 2D touch points (similar to 2D fiducial markers), or modulated touch frequency. Chan et al. showed CapStones and ZebraWidgets [2], which supports stackable tangible

Figure 3: Touch patterns for capacitive multi-touch displays.

objects on mobile devices like iPads. Voelker et al. [16] presented research on creating stable touch patterns on capacitive multi-touch displays. Common features in these projects are that they all use physical objects as purely input devices/controllers and graphical output is only displayed on the tablet device beneath. Other researchers applied mobile devices in the table-top environment as smart tangible objects, to achieve various effects, e.g., magic lens [8, 12, 14] and file transfer [18, 9]. These approaches fit the ubiquitous table-top scenarios very well. However they all require certain specialized (optical/ultrasonic) tracking methods to localize mobile devices which limit their mobility and accessibility.

Inspired by these previous work, we combine 2D touch pattern tracking technique with smart devices, which extends TUIs with inexpensive nonspecialized hardware for easy deployment. Furthermore, the diverse configurations of touch patterns offer novel and flexible graphical interactions.

System Description

System overview

The proposed system has a host-client structure, with devices connected via wireless network. The host device defines a global coordinate system (CS) for all mobile displays. To detect device identities and poses, we attach 2D touch patterns (see Figure 3). Every individual device (host/client) can detect the devices stacked on its multi-touch screen and track their pose w.r.t. the global CS. The tracked view transformation is then streamed to the corresponding client devices to synchronize the display content or adjust the viewport. In addition, the host device collects gesture interactions, like panning, pinching or rotating, from all involved devices. The resulting model transformation is updated and broadcasted to all devices.

Touch patterns

Similar to [16], the touch pattern is designed using three touch points arranged in a 2D plane (see Figure 3). Touch points are in round shape with 8-9 mm diameter and 3 mm height, connected with copper foil, in order to trigger the multi-touch sensors. Distance between touch points varies from 30 mm to 70 mm. For a stable multi-touch detection, we wrap the foil around the sides of the device, so that they always connect to the user's finger (ground) during user interaction. For easy deployment and handling, a touch pattern is attached to the protecting case of a mobile device. The pairing between a physical pattern and its screen coordinates has to be calibrated only once for each device type (e.g. iPhone4 and iPhone5). The most convenient way to perform such calibration is to place a touch pattern on its device's multi-touch screen and align them physically (see Figure 3). The touch pattern coordinates are then stored locally in that client device. When it joins other devices, the coordinates are sent to the host for identification.

Device Tracking

When we place a client C on a host H's display, the system detects the client's identity and pose by comparing the current touch points P_h (in host screen coordinates) to the pre-calibrated touch pattern P_c (in client screen coordinates), see Figure 4. In our current prototype, we simply encode the client's identity in the ratio of longest to shortest edge of the triangle formed by the three touch points. Let T_c be the 3×3 matrix that has the one-time calibrated coordinates of the three 2D points in P_c in its columns and 1s in the third row and let T_h be the analogue 3×3 matrix for P_h which is updated in every frame. Then $T = T_h * T_c^{-1}$ is the 2D transform (in homogeneous coordinates) that maps pixels from the host display to the corresponding pixels on the client display. If the host device is a client to another

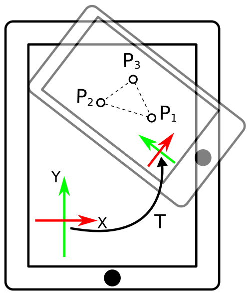

Figure 4: View transformation of magic lens.

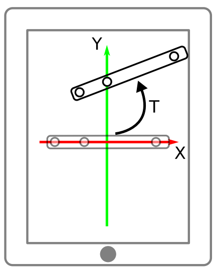

Figure 5: View transformation of bench viewer.

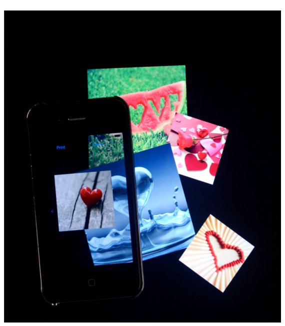

Figure 6: A photo collage can be created using the intuitive copy-paste action offered by ACTUI.

host (in a hierarchical setting) then the respective viewport transforms have to be concatenated.

Exemplary applications

We present two example applications which show potential usage of ACTUIs. By using commodity devices, ACTUIs enable diverse configurations in different scenarios. Moreover, all these applications benefit from the fact that ACTUIs are active interaction devices that can display their own visual content and thus provide a very intuitive link between the physical prop and virtual content.

Magic lens

Once we have the pose of the client device relative to the coordinate system defined by the host, we can synchronize the visual content on both displays to achieve a magic lens effect [8, 15], as shown in Figure 1. Since the client display is overlaid on the host display, by showing an extra layer of visual information on the client, we actually combine the physical property (layers of displays) with the digital content (layers of visual data). Such kind of layer display could be used in various application scenarios, like medical image visualization, augmented website browsing, or mobile gaming. Moreover, the design of our system theoretically supports arbitrary number of stacked layers, which provides even more application possibilities, e.g. filter glasses. A similar idea was proposed in [5], where several co-located mobile devices are stacked together to synchronize calendars. In addition, our ACTUI allows users to keep the visual content even when taking the client device off the host display, which enables an intuitive copy-paste operation for visual content when moving the client from one host to another. For example, photos can be copied from a client to the host and freely laid on the host screen to compose a collage, see Figure 6. A similar operation was presented in [13], where a table-top environment is required. Compared to that approach, our ACTUI is more suitable for mobile scenarios.

Bench viewer

By attaching a touch pattern to the edge of a mobile device, we can provide additional degrees of freedom for interaction, see Figure 2, 5. In this setup the client display adds another visualization dimension to the planar display surface of the host device. From an interaction point of view, the orthogonal displays provide a tangible materialization and control of the host display's third dimension and offer another viewing perspective "in situ" on the planar display. Figure 7 shows 4 types of interaction provided by this setup. The upright screen can be moved and rotated freely on the horizontal display surface to control the viewpoint in the 3D space, while panning and pinching gestures can translate and scale the virtual object. We show a demo application which visualizes a 3D MRI volume data via the ACTUI interface. The host displays a horizontal slice through the dataset for reference. When moving the client device, the corresponding vertical slice is shown on its display. The host display provides the spatial reference such that anatomic structures that extend across several slices can be inspected and traced in a very intuitive fashion. Moreover, finger gestures on the upright client display can be fed back to the host to change the horizontal slice as well. There are many more scenarios that could benefit from the extra dimension of interaction provided by the bench viewer. For example, in a scenario of photo/music 3D browsing, we can switch albums by moving the vertical screen on the horizontal surface. The content of an album can be checked by panning on the vertical screen. Or in a scenario of video editing, we can fast-forward by moving the upright screen across the surface below.

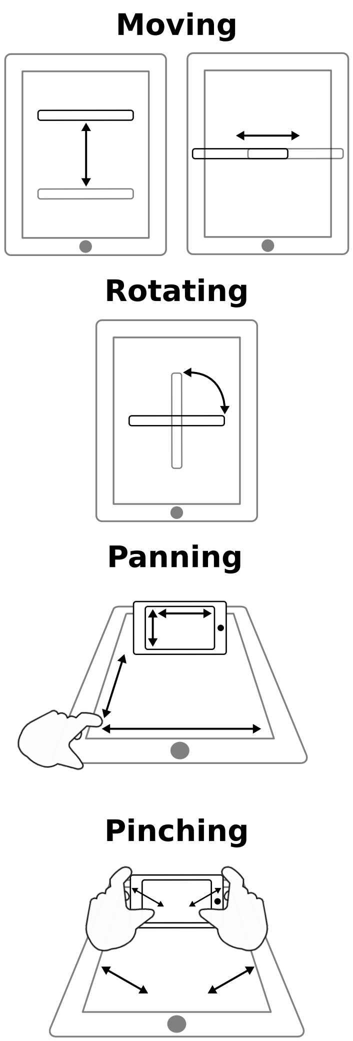

Figure 7: Possible interactions in Bench viewer.

Evaluation and Discussion

In order to evaluate the precision of the touch pattern tracking, we compared ACTUI to the ARTTrack system [1], an infrared light optical tracking system which achieves tracking precision at mm level. We put an iPad3 (host) at the origin of the optical tracking system and align the coordinate system (CS) of the iPad with the tracking system's CS. An optical marker was attached to the center of an iPhone case equipped with a touch pattern on the back. In this way we can compare the position computed by the touch pattern tracking with a ground truth, the position provided by the optical tracking system. Both positions were collected in the iPad when we moved the case on the iPad screen. We collected 8000 samples and computed the distance between each pair of positions. An average distance of 1.92 mm was found (SD = 1.86). Furthermore we measured an average angular error of 2.2 degree (SD: 1.6 degree) which, given the size of a smartphone, matches the positional precision. The main reason for the position error is that a touch point is actually not a point (8 - 9 mm diameter), so depending on the pressure on the device, the detected transformation can vary. However, we should notice that the position error is relatively small compared to the border thickness of mobile devices, e.g. iPad border 19 mm, iPhone4 4.5 mm, such that slight visual discontinuities are not noticed by users. Nevertheless, the influence of visual discontinuity to usability in this scenario should be investigated in the future.

We also measure the time consumption per touch pattern tracking (duration from the beginning of a touch event to the end of transformation update), which is less than 0.1 ms (on a 3rd generation iPad). Although the pattern tracking is very fast, there is a perceivable latency between a user action and the screen update. It mostly emerges from several system effects that our implementation cannot fully control, like touch screen sensor delay, display update rate and WiFi network latency. These effects clearly dominate the perceived latency but cannot easily be avoided due to technical restrictions.

In the current prototype, we encode the pattern ID in the ratio of longest to shortest edge of the touch pattern, which was sufficient if only a few devices are used. To increase the number of IDs, we can further encode information into all three edges. A marker with four touch points can also be applied [19].

Conclusion

We presented ACTUI, an active commodity tangible user interface that builds tangible user interfaces using simply enhanced mobile devices. By attaching 2D touch patterns on the back of mobile devices, we can track their position on capacitive multi-touch screens and synchronize their visual content. Diverse configurations of touch patterns enables novel and flexible graphical interaction. Two example applications were demonstrated to show the potential of the ACTUI concept.

Acknowledgments

This research was funded by the European Research Council (ERC Advanced Grant "ACROSS", grant agreement 340884), and the German Research Foundation (DFG, Gottfried-Wilhelm-Leibniz Programm).

References

[1] ART. Arttrack system. http://www.ar-tracking.com/.
[2] Chan, L., Müller, S., Roudaut, A., and Baudisch, P. Capstones and zebrawidgets: sensing stacks of building blocks, dials and sliders on capacitive touch screens. In *Proc. CHI '12* (2012), 2189–2192.
[3] Hilliges, O., Izadi, S., Wilson, A. D., Hodges, S.,

Garcia-Mendoza, A., and Butz, A. Interactions in the air: Adding further depth to interactive tabletops. In *Proc. UIST '09* (2009), 139–148.

[4] Kaltenbrunner, M., and Bencina, R. reactivision: a computer-vision framework for table-based tangible interaction. In *Proc. TEI '07* (2007), 69–74.

[5] Lucero, A., Keränen, J., and Jokela, T. Social and spatial interactions: Shared co-located mobile phone use. In *CHI EA '10* (2010), 3223–3228.

[6] Merrill, D., Kalanithi, J., and Maes, P. Siftables: towards sensor network user interfaces. In *Proc. Tei '07* (2007), 75–78.

[7] Microsoft. Microsoft pixelsense. http://www.microsoft.com/en-us/pixelsense/, 2011.

[8] Olwal, A., and Feiner, S. Spatially aware handhelds for high-precision tangible interaction with large displays. In *Proc. TEI '09* (2009), 181–188.

[9] Olwal, A., and Wilson, A. D. Surfacefusion: Unobtrusive tracking of everyday objects in tangible user interfaces. In *Proc. GI '08* (2008), 235–242.

[10] Pla, P., and Maes, P. Display blocks: Cubic displays for multi-perspective visualization. In *CHI EA '12* (2012), 2015–2020.

[11] Rekimoto, J. Smartskin: An infrastructure for freehand manipulation on interactive surfaces. In *Proc. CHI '02* (2002), 113–120.

[12] Sanneblad, J., and Holmquist, L. E. Ubiquitous graphics: Combining hand-held and wall-size displays to interact with large images. In *Proc. AVI '06* (2006), 373–377.

[13] Spindler, M., Büschel, W., and Dachselt, R. Use your head: Tangible windows for 3d information spaces in a tabletop environment. In *Proc. ITS '12* (2012), 245–254.

[14] Spindler, M., Bschel, W., Winkler, C., and Dachselt, R. Tangible displays for the masses: spatial interaction with handheld displays by using consumer depth cameras. *Personal and Ubiquitous Computing* (2013), 1–13.

[15] Spindler, M., Stellmach, S., and Dachselt, R. Paperlens: Advanced magic lens interaction above the tabletop. In *Proc. ITS '09* (2009), 69–76.

[16] Voelker, S., Nakajima, K., Thoresen, C., Itoh, Y., Øvergård, K. I., and Borchers, J. Pucs: Detecting transparent, passive untouched capacitive widgets on unmodified multi-touch displays. In *UIST '13 Adjunct* (2013), 1–2.

[17] Weiss, M., Wagner, J., Jansen, Y., Jennings, R., Khoshabeh, R., Hollan, J. D., and Borchers, J. Slap widgets: Bridging the gap between virtual and physical controls on tabletops. In *Proc. CHI '09* (2009), 481–490.

[18] Wilson, A. D., and Sarin, R. Bluetable: Connecting wireless mobile devices on interactive surfaces using vision-based handshaking. In *Proc. GI '07* (2007), 119–125.

[19] Yu, N.-H., Chan, L.-W., Lau, S. Y., Tsai, S.-S., Hsiao, I.-C., Tsai, D.-J., Hsiao, F.-I., Cheng, L.-P., Chen, M., Huang, P., and Hung, Y.-P. Tuic: enabling tangible interaction on capacitive multi-touch displays. In *Proc. CHI '11* (2011), 2995–3004.