Paper to be presented at CHI 2007 workshop on Tangible User Interface in Context and Theory, San Jose CA, 28 April 2007.

Tangible Interaction in a Mobile Context

Alan F. Blackwell, Gareth Bailey, Ignas Budvytis, Vincent Chen, Luke Church,
Lorisa Dubuc, Darren Edge, Mattias Linnap, Vilius Naudziunas, Hugh Warrington

University of Cambridge Computer Laboratory
William Gates Building, J J Thomson Avenue, Cambridge UK.
{afb21, gpb29, ib255, sc455, lec40, lmd37, dke22, ml421, vn229, hjmw2}@cam.ac.uk

ABSTRACT

We describe three design experiments, implementing interactive systems that explore the technical context of mobile device usage as a potential design target for tangible interaction techniques. These systems are all implemented using commodity hardware, with camera input to locate tangible interaction elements, and Bluetooth to coordinate multiple devices.

Author Keywords

Tangible User Interfaces, Mobile User Interfaces

ACM Classification Keywords

H.5.2 User Interfaces – Input devices and strategies, Interaction styles

INTRODUCTION

This paper investigates technical opportunities for tangible computing within a particular class of socio-technical contexts, based on current trends in personal computer manufacture and usage. Most tangible user interface (TUI) prototypes have until now been motivated by laboratory or office settings. The tangible elements of such interfaces are often designed to rest on a desk or table, or to be fastened to a wall. The sensing infrastructure is often immobile and/or rigidly fixed to desks tables and walls. Physical size, power consumption and processor/memory performance of the computer equipment controlling the TUI is seldom a problem. The computer can be as large as necessary, located in the same laboratory or office and connected to the sensing infrastructure via fixed or umbilical cables.

Our research starts from the observation that this technical context, although familiar in research laboratories from the earliest predecessor scenarios of experimental multimedia and multimodal user interfaces (e.g. Sketchpad and SDMS [2]), is increasingly unrepresentative of most people's experience of computer use. The most common multimedia computing devices are now mobile phones, music players and video players. Even keyboard-using knowledge workers often prefer laptop to desk-bound computers, because of the flexibility afforded by continuing to work in various locations, having wireless access to information resources at all times, and continuing work in airport lounges, hotels or meeting rooms. For textual tasks, keyboards themselves are becoming more compact, folding under, behind or over display surfaces, or even packed separately for occasional use with a pointer or keypad driven device.

Figure 1: a) N80 smartphone and b) N770 internet tablet, shown approximately to scale, with standard credit card for comparison.

This is the context for which we are designing. The devices commercially available for these environments often have powerful multimedia processing capabilities. Yet users may prefer not to think of them as computers. The two products that we have taken as typical target platforms for our research are the Nokia 770 "Internet Tablet" and the Nokia N80 smartphone. Both run general purpose operating systems (Debian Linux and Symbian respectively). Both are pocket-sized (coat pocket), but more importantly are so lightweight and low-power that they can easily be carried in a briefcase, rucksack or handbag alongside papers, books and other everyday business supplies or travel items. As consumer items, both are relatively low-priced by comparison to standard laptops, so economical alternatives. (The first author uses an N770, along with a folding Bluetooth keyboard, as his only computer when travelling).

Mobile vs. Wearable Tangible Interaction

Although the small size and low power requirements of these devices might suggest that they are ideal candidates for wearable computing, it is important to note that we are not investigating wearable computing in this research. Although there are clearly shared research concerns between tangible interaction and wearable computing, the presumed context of use is quite different. Wearable computers are designed for interaction while standing and walking, meaning that they have restricted visual channels (unless accompanied by head mounted display), and do not generally support two-handed operation (being operated in a pocket, or via controls mounted on the body). Our concern is with the travelling or casual computer user who is temporarily seated. Such a user might naturally take a device from a pocket, place it on any convenient surface (airplane tray table, hotel bedspread), and then interact directly for a period of time, watching the display and able to use both hands.

TECHNICAL APPROACH

Because our research focus is motivated by the established markets for a class of consumer devices, we wished to provide our design experiments with external validity by limiting ourselves to the I/O channels supported by this class of devices. The resources available to us were as follows. The N770 has a 7cm high resolution touch screen that supports stylus interaction, a few navigation buttons, and support for Bluetooth peripherals. The N80 has a 3MegaPixel camera, phone and navigation buttons, a "shutter" button and Bluetooth support. Both devices also support WiFi, although we concentrated on Bluetooth for these experiments. The manufacturer's expectation for these devices is apparently that a customer might buy both, and that they would be complementary (the N770 has neither a SIM card nor camera, and its user interface places Bluetooth connection to a phone as one of the foreground configuration tasks).

Our principal technical strategy was to support a range of interaction styles by exploiting the camera in the N80 as a general purpose sensing and identification channel. We used three different technical approaches that are relatively well established in the computer vision literature, but not yet commonly implemented in interactive scenarios employing low-power commodity hardware of this kind.

- The first of these is feature-based recognition of previously seen objects, in order to identify a particular object in the visual image from among a set of alternative candidates.
- The second is the use of optical flow analysis to estimate motion of the camera.
- The third is template matching of hue-based image regions to determine the position, orientation and articulation of a distinctively coloured object.

Combinations of these three techniques might allow a sophisticated application to infer a great deal about the full physical context of a camera-equipped mobile phone, including gestures made by the user, other objects the user is carrying or manipulating, and so on. However, this is probably beyond the capability of this generation of phones. Instead, our design experiments investigated one technique at a time, in order to study applications that might be feasible in our scenario contexts. We should emphasise that these are design experiments motivated by a specific context, not contextual studies of use – although we certainly hope to carry out such studies in future.

EXPERIMENT 1: LINKING MOBILE DEVICES TO EVERYDAY TANGIBLE SURFACES

Figure 2 – Bimanual interaction between camera-phone and visual tags on printed leaflet (result of previous project with Intel Research)

Handheld devices this small can be considered as tangible interaction elements in themselves, as if a whole computer had been embedded inside a mouse. The size, form and controls on the case afford certain kinds of interaction in themselves. We were particularly taken with the fact that the N80 phone, if turned on its side, appears much like a digital camera, having a shutter button on the top, a viewfinder, and a preview screen. In earlier research [6], we had used cameraphones as interaction devices, connected by Bluetooth to a network application, which were used to recognise and link from visual tags (a kind of circular barcode suitable for low-resolution) on a printed brochure.

That earlier combination of brochure-plus-phone, when interpreted as image-plus-camera suggested a bimanual interaction style in which images such as photographs that were meaningful to the user might be "re-connected" to the camera that took them, but without any obvious technical apparatus. Rather than visual tags, we used a fast algorithm that matches images based on a tree of low-level "keypoints" that are particularly salient [5]. The algorithm is insensitive to small degrees of image masking, tilt, orientation and lighting conditions, so that an image considered by the user to be unique should be recognised as such when held in front of the camera. We use this camera plus photograph interaction style to provide clear interaction cues to users with limited technical confidence (people with Alzheimer's disease), based on the physical form and affordances of familiar objects that happen to support computational augmentation when used with our system [3].

EXPERIMENT 2: BIMANUAL INTERACTION WITH MOBILE DEVICES

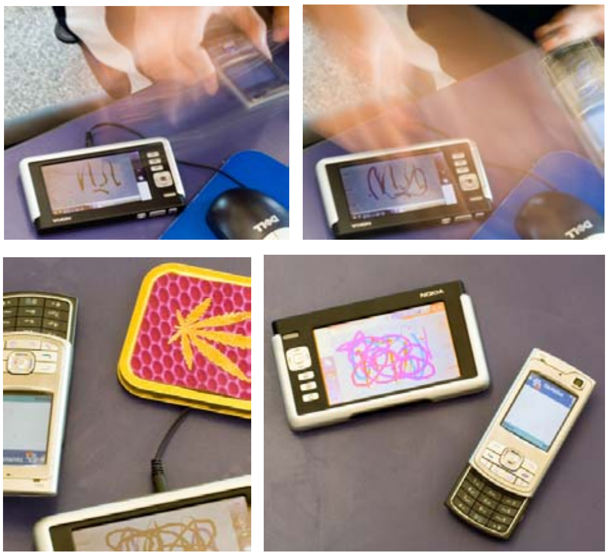

Figure 3 – a) Selecting a colour (from blue mat) by "scrubbing" gesture; b) varying brush width during bimanual interaction by "tilt" gesture; c) selection of alternative colour; d) painted result.

In this experiment, we explored bimanual interaction in which, once again, the whole device can be held in the hand as a tangible manipulation device. We imagined a user sitting "painting" with their mobile devices, working with available colours from the objects and scenery around them, (perhaps to supplement or alter a photograph of that scene).

We used an optical flow algorithm (a re-implementation of Wang and Canny's TinyMotion [7]) to estimate continuous motion of the N80 phone, so that it could be used as a supplementary pointing device. Our initial plan was to use dual pointers, as in previous bimanual interaction research, a possibility which we explored using a Bluetooth mouse in conjunction with the N770 stylus. Unfortunately, although there are multi-pointer extensions for the Linux X-window server code, these are incompatible with the N770 version of Linux. For those interested, this experiment is documented, with source code [1]. Instead, we developed an interaction style that exploited the natural affordances of the compact camera-like form factor to create a novel painting application, as seen in Figure 3. Moving the camera left-to-right (users might think of this as a "scrubbing" gesture) picks up a paint colour from the camera's visual field. Tilting the camera provides continuous variation of the brush width. If the camera is held in the non-dominant hand, and the N770 stylus in the dominant hand, the effect is to provide interactive functionality at least equivalent to the N770's native painting application, but requiring no onscreen controls. Furthermore, both colour and brush width are continuously and dynamically variable, allowing creative effects that are not possible when a single stylus alternates between paint and control tasks.

EXPERIMENT 3: INTERACTION WITH SMALL ARTICULATED TANGIBLES

Figure 4 - a) articulated tangible device; b) selection of image warp mode; c) warped image

In this final experiment, we imagined that users might make use of other objects in their mobile computing environment as tangible interaction aids. It was clear from our first experiments that camera interaction allows the use of tangible objects that have no electronic augmentation, but are simply identified and tracked from visual input. As we had previously explored the potential of generic solid object tokens and bimanual interaction under visual tracking [4], we were interested in objects that might be articulated (i.e. having internal joints), in addition to having position and orientation control. Figure 4 shows an application in which a simple articulated object is recognised by a camera, and mapped into the plane of an image that is being edited. This might, for example, be an image that has just been captured using a camera phone, but which the user wishes to crop, annotate or enhance before sending it by MMS.

We found that the use of an articulated tangible allows rapid and intuitive selection of salient features such as profiles within the image. In our demonstration application, the same tangible can then be used to specify an image warp relative to that profile. Typical entertainment applications that we discovered included the adjustment of facial expressions and body shapes. Unfortunately, the computational requirements of the image warping algorithm made local execution on the N770 impractically slow, although we are confident that the N80 could carry out the necessary visual template matching to recognize in real time the position and configuration of the tangible. A test implementation, using a standard PC with webcam input to process both camera input and image warp on the same machine, worked reliably at interactive speed (figure 5), and we believe that similar applications will very soon be feasible (say within a year) in our target class of mobile devices.

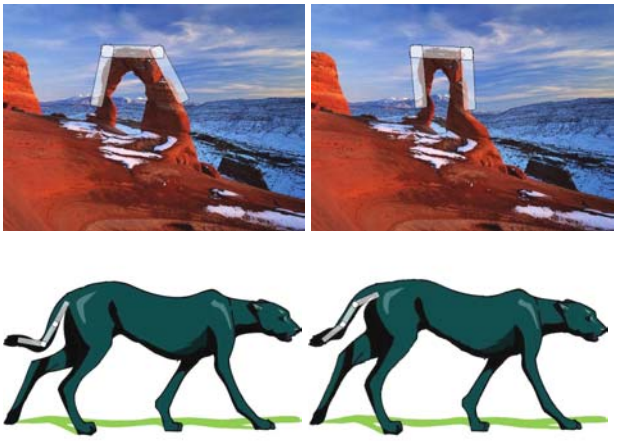

Figure 5 – Images before (a and c) and after (b and d) profile warp operations executed on a standard desktop computer.

Our demonstrator application, controlling amusing warping of an image, appears to have plausible appeal to the target market (cameraphone users who might make amusing adjustments to an image when temporarily seated with mobile technology to hand). However the main research objective here was to explore the ways in which tangible interaction in this context might provide richer use of the hands and fingers than simple pointing, object selection, and motion gestures. The articulated object that we made has five continuous degrees of freedom (X-Y position on the table, rotation, and the angle of the two internal joints). It can easily be manipulated with one hand, and its potential correspondence to salient image features is directly and immediately understood by users. We believe that these features offer powerful opportunities for new tangible interaction devices that are feasible even in this constrained context.

IMPLICATIONS FOR DESIGN

We have demonstrated a range of new interaction opportunities, using commodity hardware that is typically used in contexts very different from those of conventional TUI research. Rather than large scale sensing surfaces and interactive environments, we have focused on very small scale devices of a size that might easily be carried in the user's pockets, and laid out for use on an airplane tray table. We believe that this is a rich context for the application of TUI interaction principles, especially given the potential of mobile device casings to be handheld tangible objects in themselves, and the opportunity for small articulated objects to be recognized and tracked visually. These simple objects, in conjunction with camera input, can integrate mobile interaction with a variety of physical objects in the user's environment, as a tangible supplement to existing touchscreen, button and stylus interaction.

ACKNOWLEDGMENTS

This research was supported by Boeing Corporation, by Nokia, and by the Engineering and Physical Sciences Research Council undergraduate bursary programme. The first experiment was implemented by Mattias Linnap, with Luke Church and Lorisa Dubuc. The second experiment was implemented by Gareth Bailey, Vincent Chen and Hugh Warrington. The third experiment was implemented by Ignas Budvytis and Vilius Naudziunas with Darren Edge.

REFERENCES

1. Bailey, G. (2006). Bluetooth Mice. Contribution to maemo-developers mailing list, 17 August. http://maemo.org/pipermail/maemo-developers/2006-August/005153.html
2. Blackwell, A.F. (2006). The reification of metaphor as a design tool. ACM Transactions on Computer-Human Interaction (TOCHI), 13(4), 490-530.
3. Dubuc, L. (2006) An Accessible Technology to Improve Communication in Dementia. In Proceedings of IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC), pp. 226-227.
4. Edge, D. (2006). A Tangible User Interface for Peripheral Task Management. In Proc. IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC), pp. 228-229.
5. Lepetit, V. and Fua P. (2006). Keypoint Recognition Using Randomized Trees. IEEE Trans. Pattern Anal. Mach. Intell. 28(9): 1465-1479.
6. Toye, E., Sharp, R., Madhavapeddy, A., Scott, D., Upton, E., and Blackwell, A. (2006). Interacting with mobile services: An evaluation of camera-phones and visual tags. Personal and Ubiquitous Computing, Jan 2006, pp. 1 - 10
7. Wang, J. and Canny, J. (2006). TinyMotion: Camera Phone Based Interaction Methods, in alt.chi of ACM CHI 2006