Mixed Interaction Space
– Designing for Camera Based Interaction with Mobile Devices

Thomas Riisgaard Hansen
Center for Pervasive Healthcare
Department of Computer Science,
University of Aarhus
Aabogade 34, DK8200
Aarhus N, Denmark
thomasr@daimi.au.dk

Eva Eriksson
Interactive Spaces, ISIS Katrinebjerg
Department of Computer Science,
University of Aarhus
Aabogade 34, DK8200
Aarhus N, Denmark
evae@interactivespaces.net

Andreas Lykke-Olesen
Department of Design
Aarhus School of Architecture,
Nørreport 20, DK8000
Aarhus C, Denmark
alo@interactivespaces.net
+4589360000

ABSTRACT
In mobile devices, such as mobile phones and PDAs, an integrated camera can be used to interact with the device in new ways. In this paper we introduce the term mixed interaction space and argue that the possibility of using mixed interaction spaces is what distinguishes camera-based interaction from other types of sensor-based interaction on mobile devices. We present our implemented applications, and related work that use mixed interaction spaces. Based on this we address how mixed interaction spaces can have different identities, be mapped to applications, and how it can be visualized.

Author Keywords
Input and interaction technologies, augmented reality, tangible UI, interaction design.

ACM Classification Keywords
H5.2. User Interfaces (*Interaction styles, Haptic I/O, GUI*)

INTRODUCTION
An increasing amount of today's mobile devices are equipped with integrated cameras, which can be used to determine how devices are manipulated. By applying image analysis algorithms on the camera pictures, the movement, and in some cases the rotation and tilting, can be determined. This input technology has been used to implement a set of different applications by e.g. SpotCode [13], SemaCode [12] and Rohs [10]. To a great extent the focus in these projects is on the technology itself whereas the interaction technique is not discussed or analyzed.

In this paper we introduce the term mixed interaction space and argue that the possibility of using the position in space distinguishes interaction techniques based on the integrated camera from other interaction techniques that e.g. use accelerometers or compasses as data input sensors. We present our own work with mixed interaction spaces and argue how it relates to other projects that use similar techniques. We then discuss mixed interaction spaces in detail and point out three important characteristics; identity, mapping and visualization.

CHI 2005, April 2–7, 2005, Portland, Oregon, USA.

ACM 1-59593-002-7/05/0004.

MIXED INTERACTION SPACE
There exist several novel interaction techniques for mobile devices to supplement button and pen interaction. Speech is an obvious candidate; the speech recognition systems available on mobile devices can efficiently be used to select a specific command e.g. calling a specific number. Voice commands, even for simple navigation introduce cognitive overhead [4], which can be a problem.

Accelerometers, sometimes combined with a compass, can interact with an application by using tilting, rotation and movement of the device as input. The clear advantage of this interaction technique is its independence of the surroundings why it supports mobility very well. It supports new ways of interacting with applications e.g. scrolling in applications by tilting the device [6].

Interaction techniques that use integrated cameras strongly resemble interactions that can be designed with accelerometers. The movement, rotation and tilting of the device, can partly be extracted from running optical flow algorithms on the camera images. However, the camera images can provide more information than the movement, tilting or rotation vector. It can be used to identify a feature, or fixed point, and it can calculate its relative rotation, tilting and position according to this point. A space is spanned from this fixed point to the end of the camera view (see Figure 1), and this space is what we call *the mixed interaction space*.

With *mixed* we try to emphasize that the space is a *physical space*, but at the same time the space plays an important role in the *digital interaction* that is controlled by the movement of the device in the space.

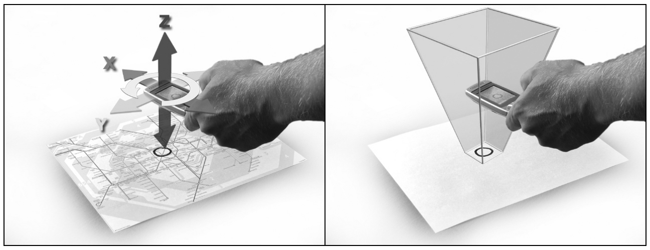

Figure 1: Interaction gestures and a diagram of the mixed interaction space.

The mixed interaction space has the shape of an inverted pyramid (see Figure 1). From the tracked fixed point, a space is spanned bound by the camera's ability to see and detect the fixed-point. When the device is close to the fixed-point the orthogonal movement plane in which the mobile device can track the circle is small. The size of the movement plane increases as the mobile device gets further away from the fixed point, but tracking also gets more difficult. The mixed interaction space ends when the mobile phone is unable to see or track the fixed-point. Large fixed-points will span larger mixed interaction spaces that are suitable for large movements and small fixed-points span smaller mixed interaction spaces suitable for small gestures.

Mixed interaction space can be implemented in standard mobile devices equipped with a camera, and does not require external sensor technology. The concept is built on the principles of direct manipulation [11], the actions are rapid, incremental, and reversible and the effect is visible immediately. The user is able to act through gesturing, and feedback occurs immediately which convey the sense of causality between the gesture and the application.

Related work

Several other projects have worked with using the integrated camera to interact with applications; Rohs have implemented a tracking system in mobile phones that tracks a 2D barcode tag with two guiding bars. The device calculates the rotation, tilting, identity and position of the mobile device in the 3D space [10]; The SemaCode project has made a system running on mobile phones that translate 2D barcode tags to URLs [12]. The project does however not give the position or rotation of the device in the 3D space; SpotCode is another project for mobile phones and works with circular barcode tags [13]. This system is able to give the rotation, position and identity of the tag in the mixed interaction space.

Another related project is MouseField by Masui et al. [7]. In this project the movement and identity of everyday objects like e.g. CDs are tracked and used to interact with different applications. A set of actions can be associated with each object but in contrast to mixed interaction space, only a 2D and not a 3D space is spanned, and used to remote control appliances.

IMPLEMENTATION AND APPLICATIONS

We have worked with mixed interaction spaces in several projects. Our initial approach was to implement a system that did not rely on tracking barcodes, but we wanted to track something that does not have to be created in a computer, and that does not stick out in the environment like 2D barcodes. In our project we have chosen to use a simple circle as tracking point, as the circle has several advantages. For instance, a circle can be hand drawn or it can be part of a common artifact. We explored for instance how the edge of a watch could be used as a tracking point.

To detect the fixed point, the Randomized Hough Circle Detection Algorithm as described by Xu [15] is implemented and optimized for detecting a circle in a picture. The system is implemented in C++ for Symbian OS 7.0s on a Nokia 7610 mobile phone. To keep the interaction fluent and to reduce the memory used, video is captured in a resolution of 160x120 pixels or 320x240 pixels, depending on the feedback requirement from the application.

The current implementation tracks a black non-perfect circle on a mainly non-black surface in different light conditions, and the interaction space has a stable zone in the center. The implemented algorithm is the base for four different applications.

With our tracking system we implemented the 'ImageZoomViewer' application that allows the user to pan and zoom simultaneously on a map or image. In 'Flower', we implemented an augmented layered pie menu application. When moving the phone in the mixed interaction space we browse between the different menus, and pan is used for choosing within a pie menu. In 'Drozo', we have implemented a system that allows the user to rotate, move, and scale images on a large screen with the mobile device. Similar applications have been implemented by Rohs[10], SemaCode[12], SpotCode[13].

However, because we rely on a simple feature that everyone can draw, we have implemented the application DrawME that allows the user to draw a circle with a symbol in the center, and associate this symbol to a specific interaction. We have for instance worked with connecting special phone numbers to specific symbols. With the DrawME application, as soon as a symbol is recognized, a pie menu appears on the mobile phone asking the user if s/he wants to call the contact associated with the symbol. When moving the phone to the right in the mixed interaction space this person will be called and moving the phone to the left the menu will be closed.

In the projects we have worked with we have found that mixed interaction space provides a novel way of looking at interaction with mobile devices. Button pressing and pen interaction might still be the most efficient input methods for most applications running on mobile devices, however we have found that mixed interaction spaces open up for novel applications that, to a larger extent use the physical space around the mobile device, that is unused by button and pen input applications.

DISCUSSION

Based on our work with mixed interaction spaces we present three design issues and possible solutions, designers exploring these spaces need to address (see Table 1). The design issues are: Is it possible for mixed interaction spaces to have an *identity*; how should the *mapping* between the physical movement of the device in space and the action on the interface be; and finally, how is the interaction space *visualized*?

Design issues:	Identity	Mapping	Visualizing the interaction space
Design solutions:	Simple fixed point Identity interfaces	Natural mapping Semantic mapping	Overlays Icons Tactile feedback

Table 1: Table of design issues and solutions of the mixed interaction space.

Identity

The identity of the mixed interaction space differs depending on the type of application. We have made a distinction between interfaces based solely on a simple fixed point (simple fixed-point interfaces) and interfaces where the fixed point has a specific identity (identity interfaces).

Fixed-Point Interfaces: The simple fixed-point interface only needs the software to recognize a fixed point. The fixed point can be a circle, a hand-drawn symbol, or a personal artifact like a finger ring. Simple fixed-point interfaces can e.g. be used to navigate in applications in the device. The symbols can be seen as physical keys into the digital world, and thereby resemble tangible interfaces that aim to distribute the different controls to the real world. One of the problems with tangible interfaces, as pointed out by Greenberg [5], is that in order to use them, they require special tangible objects, which is not the case with mixed interaction spaces.

Identity Interfaces: In identity interfaces the interaction space is associated with a specific identity. The identity can be extracted by reading a barcode [13], using short range Bluetooth [2] or by RFID tags [14, 7]. With identity interfaces, a specific interface can be associated with each mixed interaction space. The corresponding interface can be stored in the interaction device, transmitted through e.g. Bluetooth or downloaded from the internet. Identity interfaces can be used for controlling external devices, or as a direct link to digital material.

Using mixed interaction space to interact through identity interfaces can be seen as a possible method to interact with the invisible computer [9]. With identity interfaces the mobile device can be the interface to an invisible system. A fixed point placed in the context can be used as a visual cue, signalizing the existence of a hidden interface, and it can be used as a fixed point for the mixed interaction space. This way, the context can be used to reduce user interface complexity.

Mapping

Mapping is a term that refers to the relationship between two objects or things. Two different types of mapping were present in the applications we explored, natural and semantic mapping.

Natural mapping: In applications with tight couplings between the physical movement and the movement in the application natural mapping is accomplished, which is a term suggested by Norman [8]. Natural mapping uses physical analogies or cultural meanings to bring about immediate understanding of the relationship between the physical and digital movement. An example of this is the application for map navigation on a device, where moving the device to the left, right, up, or down pans an image, and moving the device closer or further away from a fix point zooms in and out. For applications that control external or internal digital objects, rotation is used to rotate the current object. This resembles Norman's example of natural mapping in which turning a steering wheel to the right make a car turn right [8].

In relation to natural mapping two mapping strategies can be pursued: Absolute or relative mapping. With absolute mapping a specific position in the mixed interaction space is mapped to, for instance, a specific zoom level and position on a map. With relative mapping a specific position in the mixed interaction space is mapped to a movement vector.

We found that relative mapping is best suited in most applications. This is due to the shape of the mixed interaction space, since it has the form of an inverted pyramid (see Figure 1). This property makes mixed interaction space unsuitable for e.g. absolute mapping of a device to a specific position on a map or at least absolute positioning on all three axes. Another problem with absolute mapping is that the image captured by the camera must have a similar size to the picture being watched. If not a small movement with the device will make the picture jump several pixels. The disadvantage with relative mapping is that it does not provide the same spatial awareness as absolute mapping of the location in the interface.

Semantic mapping: The second kind of mapping identified is what we call semantic mapping. Here moving the device in a specific direction does not necessarily map to the interface moving in the same direction. With semantic mapping a metaphor is used to bridge between the physical movement and the action on the device e.g. moving the device to the left could be mapped to the action 'play media file', 'next song' or any other command [7].

A characteristic of semantic mapping is that it is discrete; the space is divided into different layers. For an application with superimposed pie menus in several layers (e.g.

'Flower'), there is no 'natural' mapping for movement along the z-axis; 'zoom-in' can equally be mapped to selecting the currently highlighted menu item, or to moving to the next layer of menus.

Visualizing the interaction space

Visualization becomes important with sensor interfaces as pointed out by Bellotti et al. [1] in their discussion on sensor interfaces. Chalmers et al [3] have raised the question of how to visualize uncertainty in input data. Visualization is also important in mixed interaction spaces since the boundary of the space, and thereby the interaction, depend on what the camera sees and not what the user sees.

In the systems we have worked with we have found that there are three different approaches to visualize the interaction space, overlays, icons, and tactile feedback.

Overlays: In the first approach camera recordings are displayed on the screen of the mobile device and the interface is superimposed on top of that camera view. This greatly helps the user in keeping the mobile device within the mixed interaction space when interacting, since the borders of the superimposed view is the border of interaction. The downside is that the screen estate of the mobile device is limited, and superimposing the user interface on top of the camera image is not suitable for all kinds of applications, as for instance map-navigation.

Icons: In the other approach different icons are displayed in the interface, and indicate where the tracked fixed-point is located in relation to the mobile device; if the application has lost track of it, if the application is zooming in or out, etc. This approach provides the user with a minimum of distraction from the main interface, but has the disadvantage that if the device loses track of the fixed-point it is difficult to guide the user back into the interaction space.

Tactile feedback: This is so far an area that has not been investigated for interaction within the mixed interaction space. Integrated hardware, such as vibrator and audio alerts within the mobile device, offers the possibility of indicating when the borders of the mixed interaction space are crossed. Tactile visualisation can be a solution to limited screen space, but may be considered as too abstract. Tactile feedback is not suitable for sensitive or detailed interaction, since a user's hand vibrates or the user might physically react with a movement to the sound of an alarm.

CONCLUSION

With the introduction of mixed interaction space we have described an interaction technique that uses the position of the mobile device in relation to a tracked point as input. We have argued that the possibility of using mixed interaction spaces is one of the things that distinguish camera-based interaction from other types of sensor-based interaction.

We have presented our work with simple fixed-points and mixed interaction spaces, and we have discussed three important design issues in relation to mixed interaction spaces: the ability to have mixed interaction spaces with identities, how the movement in the mixed interaction space can be mapped to different applications, and finally how this invisible space can be visualized.

ACKNOWLEDGMENTS

We would like to thank people at Center for Pervasive Healthcare and Center for Interactive Spaces under ISIS Katrinebjerg, especially Kaj Grønbæk and Jakob Bardram.

REFERENCES

1. Bellotti, V., Back, M., Edwards, WK., Grinter, RE., Henderson, A., Lopes, C. Making sense of sensing systems: Five questions for designers and researchers. CHI 2002. (2002), 415-422.
2. BlipNodes. http://www.blipsystems.com/
3. Chalmers, M., MacColl, I., Seamful and Seamless Design in Ubiquitous Computing, *Technical Report Equator-03-005*, 2004.
4. Christian, K., Kules, B., Schneiderman, B., Youssef, A. A comparison of voice controlled and mouse controlled web browsing. Proceedings of the fourth international ACM conference on Assistive technologies. (2000), 72—79.
5. Greenberg, S., Boyle, M., Customizable Physical Interfaces for Interacting with Conventional Applications. In CHI Letters, 4, 2(2002), 31-40.
6. Harrison, B.L., Fishkin, K.P., Gujar, A., Mochon, C., Want, R. 'Squeeze me, hold me, tilt me! An exploration of manipulative user interfaces', in ACM CHI'98, ACM Press, Los Angeles, CA (1998).
7. Masui, T., Tsukada, K., Siio, I. MouseField: A Simple and Versatile Input Device for Ubiquitous Computing, In Proceedings of UbiCom2004, Springer (2004), 319-328.
8. Norman, D. The Design of Everyday Things, Doubleday, New York, (1999), p 23.
9. Rehman, K., Stajano, F., Couloris, G. Interfacing with the invisible computer. Proceedings of NordiCHI 2002, ACM Press (2002), 213-216.
10. Rohs, M., Real-World Interaction with Camera-Phones, 2nd International Symposium on Ubiquitous Computing Systems (UCS 2004), 2004.
11. Scneiderman, B. Designing the user interface. Addison-Wesley, USA (1998).
12. SemaCode. http://semacode.org/
13. SpotCode. http://www.highenergymagic.com/spotcode/
14. Want, R., Fishkin, K., Gujar, A., Harrison B., Bridging Physical and Virtual Worlds with Electronic Tags. In Proc. CHI 1999, ACM Press (1999), 370-377.
15. Xu, L., Oja, E., Kultanen, P., A new curve detection method: Randomized Hough Transform (RHT). In Pattern Recognition Letters 11, (1990), 331-338.