Design of Affordances for Direct Manipulation of Digital Information in Ubiquitous Computing Scenarios

Lucia Terrenghi

University of Munich, Media Informatics
Amalienstrasse 17, 80333 Munich, Germany
lucia.terrenghi@ifi.lmu.de

Abstract. This work focuses on interface design for supporting users of instrumented environments to interact with virtual information embedded in the real world. A gesture-based interaction paradigm is presented and the prototype of an interface providing affordances for gesture-based direct manipulation of digital information is described.

1 Introduction

The emergence of ubiquitous computing [18] scenarios promises to bring digital information on the walls of everyday life environments, embedding computing and displaying capabilities within the physical space and the material artifacts which populate it. In such a setting, where people can move around rather then seat at a desktop behind their screens, handle physical and digital objects at the same time, interact on shared displays in co-located or remote collaboration, the WIMP interaction paradigm comes short in supporting users' interaction. More natural ways of manipulating information need to be thought through. This paper points out the new challenges that mixed reality raises in terms of interaction design, and suggests a new gesture-based direct manipulation paradigm for instrumented environments, where technology is embedded in the physical space so as to display and sense information. The prototype of a user interface based on a mug metaphor developed within the FLUIDUM project [6] and introduced in [17] is described. In addition, a reflection on the new design perspective that needs to be taken when merging digital and physical information is presented.

2 Direct Manipulation in Instrumented Environments

In the Personal Computer environment, direct manipulation describes the activity of manipulating objects and navigating through virtual spaces by exploiting users' knowledge of how they do this in the physical world [16]. The three main principles of direct manipulation are:

- continuous representation of the objects and actions of interest;
- physical actions or presses of labeled buttons instead of complex syntax;
- rapid incremental and reversible operations whose effect on the object of interest is immediately visible.

Direct manipulation is the basis for the dominant WIMP paradigm (Windows, Icons, Menu, Pointer), with which we manage different applications; according to the activities they support, applications rely on different metaphors. In the Microsoft Office software package, for instance, visual and auditory icons mimic the objects of a real physical office. In software programs for graphic design, icons resemble brushes and pencils. While the metaphor varies according to the domain (which translated to instrumented environments could be office, living room, kitchen, etc.), the general paradigm does not change. In addition, the appearance of widgets for desktop GUIs remains consistent, and suggests affordances for mouse and keyboard interaction, e.g. 3D effects for clicking buttons, white fields for text entry, ripples on the moving part of scrollbars for dragging. Although talking about direct manipulation, in the desktop environment we mostly need indirect input devices, such as mice, track pads or joysticks, to interact with the system. The desktop metaphor maps elements of the GUI to objects of the real world: those in turn provide affordances for gesture-based direct manipulation. When mixing virtual and real information the need of providing affordances for new interaction emerges (see Figure 1).

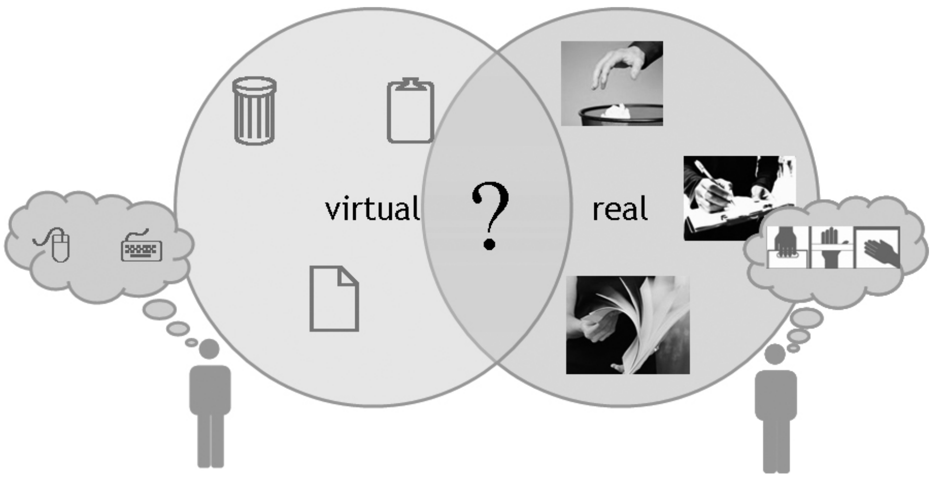

Fig. 1. In the Personal Computer environment classical GUIs rely on visual cues and metaphors in order to suggest the interaction operated with mouse and keyboard on virtual information. Real world objects provide affordances for manipulation. Affordances for information in mixed reality need to be designed for users' construction of an interaction conceptual model.

2.1 Related Work

With the emergence of ubiquitous computing scenarios, some work has been done to translate direct manipulation from the desktop environment. Butz at al. [1] try to transfer the window-desktop metaphor to 3D Augmented Reality Environments where each document type as well as each application has a visual representation in the form of a 3D icon. The operations on these icons are similar to the ones known from the window-desktop metaphor, as for example drag and drop or click to open. Other research has been focusing on devices that can work as universal remote controllers [12] or support different functions depending on the way they are physically manipulated [13][2]. Other approaches look at natural input techniques such as gesture and gaze input [19], or to tangible user interfaces that work as tokens of information to be manipulated in the physical space [7].

In this paper we present a novel interaction paradigm, aiming at direct manipulation of units of information across different displays and contexts, avoiding the use of mouse and additional control devices. In such a paradigm, surfaces act as interfaces, and hands as control devices. Ringel et al. [15] have worked in a similar direction, looking at direct hands-on manipulation without implements on SMARTBoards: differently from such approach, our attempt is not to map mouse-based interaction to hands-based ones on a touch screen display. The mouse, indeed, has a limited manipulation vocabulary (e.g. click, double click, click and drag, right click) while hands, fingers and gestures provide a much more varied one (e.g. press, draw a circle, point, rotate, grasp, wipe, etc.). Rekimoto [14] exploits such variety working on a two-hands, multiple fingers gestures vocabulary. The limit of such work is that the user has to rely on the memory of a set of actions to be performed, in order to operate with the system: the memory of such set of actions is not supported by an explicit mapping between perception and action, which is the essence of affordances.

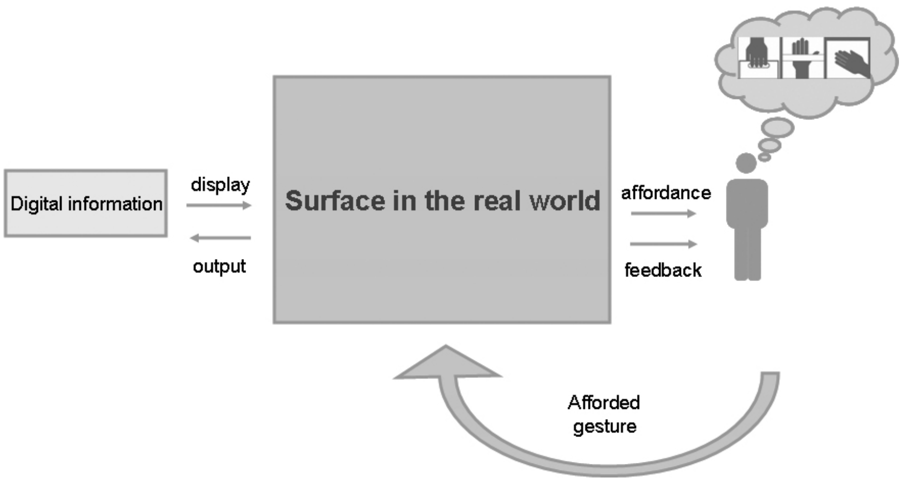

Fig. 2. The representation of abstract digital information should present affordances for gesture-based manipulation when appearing on the surface. In such a paradigm, surfaces play as interfaces and hands as control tools.

Our intent, therefore, is to design affordances for the representation of digital information which can suggest the hand gestures to be performed by the user (see Figure 2). Muscular memory can be used to structure frequent, patterned tasks into stylized or abbreviated gestures, thus providing the possibility to build a gesture vocabulary that exploits humans' sensorymotor perception [9] and haptic memory.

3 Design of an "Affordable" Interaction Metaphor

Objects of everyday life already carry information per se: their shape, color, texture, weight, temperature, material, all the aspects related to their physical features. Ecological approaches [5] focus on perception and action based on human attributes: in this context affordances can be described as properties of the world defined with respect to people's interaction with it [4]. We perceive the environment directly in terms of its potential for action: in the physical experience we explore the outside world with our senses, making inferences of objects physical behavior thanks to a semantic of perception. This is the main source for the creation of users' conceptual models of how things work. The concept of affordances is familiar especially to product design: Norman [10] applies the concept of affordances to every day life artifacts.

While designing affordances for digital information, two main design aspects can be addressed. On the one hand the visual appearance of the displayed information can suggest the gesture to be performed just relying on shapes and visual cues. Frames define semantic areas of interaction on the screen. 3D effects bring the information in fore- or back- ground and can suggest pressure. These visual cues can enhance perceptual and cognitive viewers' processing of information [11]. When objects' pliancy, i.e. their characteristic to be interactive [3], is hinted by the graphics, people get an idea of the action to be taken. In this case marks are, in semiotic terms, *nomically* related to their referents, rather then metaphorically, thus do not require viewers' interpretation of symbols [4].

On the other hand the metaphoric link to real world objects and to their affordances in the physical world can provide rich material for the design of affordances for digital information. The virtual representation of a lever doorhandle, for instance, can be mapped to the natural gesture we do when we exercise some pressure on a real physical lever. The representation of a steering wheel can be mapped to the turning gesture we perform while driving. In this case the mapping relies on the analogy between the elements representing digital information, and the affordances of their referents in the physical space.

Metaphors have long been used in GUIs for providing an intuition of how things work using the world's knowledge. While the desktop metaphor suits the type of environment in which the computing capabilities have been mostly applied so far, it runs short in scenarios of ubiquitous computing. Furthermore the visual affordances of the metaphoric items (e.g. folders and 2D icons) are suitable for the mouse-based manipulation vocabulary, but not for a hands-based one. Real world objects have affordances for manipulation and are embedded in conceptual models: digital representations of real world objects can rely on similar affordances and similar conceptual models. Building on these assumptions we have been working on the design of an interaction metaphor that

- is affordable for hands-based manipulation
- suits different environments.

4 Affordances Design Issues

In order to accomplish the direct manipulation principle of continuous representation of objects and actions, information needs to be consistently represented across a variety of contexts and to provide feedback. This helps to build a mental model: we form hypothesis of system function and physically probe and test them with our action. A preconceived model influences both the final impression and the manner in which we explore. Additionally, information may appear differently when associated with different objects, or when assuming a different status. In this sense the representation of virtual information should provide affordances that can be mapped to a certain status and suggest certain actions.

A main aspect of affordances is that physical attributes of the thing to be acted upon are compatible with those of the actor [4]: thus, representing digital information in an "affordable" way means to consider ergonomic aspects such as dominant hands, hands size, users' height, and so on.

In the paradigm illustrated above, which describes surfaces as interfaces and hands as controls, the main differences between hands and mice as operating tools need to be taken into account. A first simple difference is that hands allow for multiple simultaneous inputs. Reflecting on how we manipulate physical objects, we can easily notice hands cooperative work. For instance, we usually hold a glass with the non-dominant hand and pour the content of a bottle with the dominant hand; we hold a vase with the non-dominant hand and open the lid by rotating it with the dominant one.

The fact that there is no spatial distance between physical input and digital output also implies additional considerations, such as shadows and visual angles. Furthermore, while the ratio between the pointer and the display sizes remain constant in a mouse-based interaction, i.e. the pointer area displayed on a screen scales proportionally to the screen size, in a hands-based interaction the ratio varies in function of hands sizes.

5 The Mug Metaphor Interface Prototype

The mug interface explores the idea to rely on the affordances provided by a mug, and to metaphorically represent it as a container of information. When manipulating a real mug we know we can move it around by holding its handle, and incline it to pour its content (Figures 3a, 3b). Empty mugs are expected to be lighter than full ones (e.g. contain less data), smoking mugs are expected to be hot (e.g. contain recent data).

Additionally, a mug is an everyday life object which we use in different environments, e.g., in the office, in a living room, in a kitchen.

A first prototype of such a mug metaphor interface has been built in order to investigate the possibility to map the affordances of real world objects to gestures, relying on the conceptual model in which such real objects are embedded. In such a concept, mugs and units of information, the latter represented as kind of drops, can be manipulated across the display. The dominant hand, e.g. the right one, is devoted to the manipulation and navigation of information. A pie menu displaying containers of information is displayed in correspondence of the right hand (Figure 4a). The non-dominant hand (e.g. the left one) works as command invocation, managing a menu of resources (e.g. drain, displays, printers). Such menu can be scrolled with a movement of the finger on a holed gear, which makes the circle segments rotate (Figure 4b). The dominant hand moves units of information to the preferred resource (see Figure 5). The pie menus appear in correspondence of the hands, thus "following" the user while moving across the display, rather than being operable just in a fixed location on the screen. This responds to the need of freedom of movement of the user, and to enable two-hands cooperative interaction.

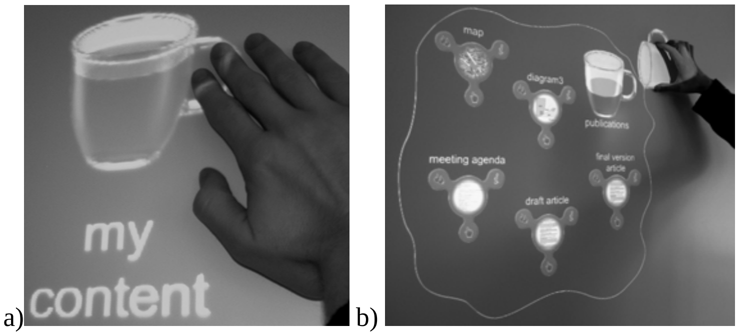

Fig. 3. The mug metaphor interface. a) To move the mug/information container, the user touches its handle and drags it on the screen surface. b) To explore its content the user turns the mug.

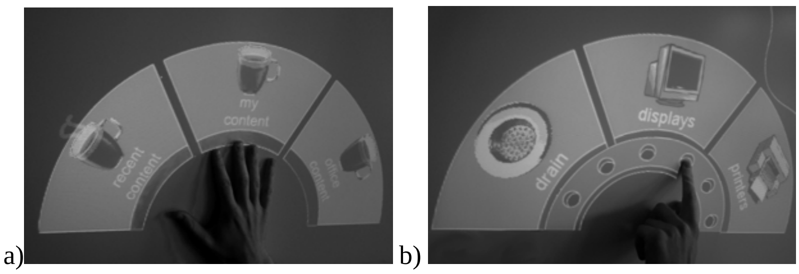

Fig. 4. a) The dominant hand is devoted to the management of information. b) The non-dominant hand is devoted to management of resources.

Fig. 5. The two hands cooperate in manipulating information: the dominant one drags a unit of information on the drain to delete it.

6 Conclusions and Discussion

In this paper an interaction paradigm for gesture-based direct manipulation of digital information in an instrumented environment was presented. In addition, the prototype of an interface based on a mug metaphor was introduced as example of ongoing research in the direction of affordances design for digital information.

6.1 Future Work

This work is still at an early stage and several issues concerning gesture recognition based on camera vision still need to be solved. Even though the paper focuses on the visual appearance of affordances, future work will explore auditory and haptic displays in order to provide redundancy and feedback. The tactile and audio channels, indeed, allow us to make sense of information in different ways. While focusing on a visual task, audio displays can influence our awareness of the environment: for instance, sound cues can draw our attention to different objects in the space (e.g. the noise of dripping water for pending tasks). Haptic displays can provide feedback to the action afforded by a visual cue: e.g., the force to be applied on a lever doorhandle can be coupled by a haptic feedback. In this sense the multimodality of interaction with digital information promises to provide more analogue interfaces.

6.2 A New Design Perspective

As emphasized in this paper, the design of affordances for digital information embedded in a real physical environment implies the consideration of new aspects which differ from the desktop PC environment. The users' possibility to move around in the space and to directly manipulate objects and information items needs to be supported by interfaces that are properly scaled to users' metrics, locations in the space, motor capabilities. The physicality of the real environment in which interaction takes place will have an impact on the interaction itself: issues such as the height of the user, her visual angle, the reachability of displayed objects to the hands, the proportion between objects and hands sizes, are some examples. In order to face such issues, ergonomic considerations need to be included in the interface design, thus suggesting the emergence of a novel design practice. The traditional usability guidelines for visual displays [8] will most likely need to be revised in order to address the novel aspects brought by ubiquitous computing. In these scenarios we expect that the design discipline will need to merge screen and product design competences, in order to merge virtual and physical worlds.

Acknowledgments. This research has been funded by Deutsche Forschungsgemeinschaft (DFG) within the FLUIDUM (Flexible User Inter- faces for Distributed Ubiquitous Machinery) project.

References

1. Butz, A., Beshers, C., Feiner, S., Of Vampire Mirrors and Privacy Lamps: Privacy Management in Multi-User Augmented Environments", ACM UIST Symposium on User Interface Software and Technology, San Fraincisco, 1998.
2. Cao, X., Balakrishnan, R., VisionWand: Interaction Techniques for Large Displays using a Passive Wand Tracked in 3D. ACM UIST Symposium on User Interface Software and Technology, 2003.
3. Cooper, A., Reimann, R., About Face 2.0: The Essentials of Interaction Design. Wiley, 17 March, 2003.
4. Gaver, W., Technology Affordances. In Proc. ACM CHI 1991.
5. Gibson, J. J., The Ecological Approach to Visual Perception. Hughton Mifflin, New York
6. FLUIDUM project, FLexible User Interfaces for Distributed Ubiquitous Machinery, http://www.fluidum.org
7. Ishii, H., Ullmer, B. Tangible Bits: towards seamless Interfaces between People, Bits, and Atoms. In Proc. CHI 1997.
8. ISO 9241-11:1998. (1998). Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability.
9. O'Regan and A. Noe, "On the Brain-basis of Visual Consciousness: A Sensory-Motor Approach", in Vision and Mind, ed. A. Noe and E. Thompson, MIT Press, 2002.
10. Norman, D.A., The Psychology f Everyday Things. Basic Books, New York, 1998.
11. Paley, W.B., Illustrative Interfaces: Building Special-Purpose Interfaces with Art Techniques and Brain Science Findings. In Proc. Smart Graphics 2003, Heidelberg, Germany.
12. Rekimoto, J. Pick and Drop: a Direct Manipulation Technique for Multiple Computer Environments. In Proc. UIST'97, (1997).
13. Rekimoto, J., Sciammarella, E.: ToooolStone: Effective Use of the Physical Manipulation Vocabularies of Input Devices. ACM UIST Symposium on User Interface Software and Technology. p. 109-117, 2000.
14. Rekimoto, J. SmartSkin: an Infrastructure for Freehand Manipulation in Interactive Surfaces, ACM Press, CHI 2002, 113-120.

15. Ringel, M., Berg, H., Jin, Y., Winograd, T. Barehands: Implement-Free Interaction with a Wall-Mounted Display. ACM CHI Conference on Human Factors in Computing Systems (Extended Abstracts) p.367-368., 2001.
16. Shneiderman, B., Direct manipulation: A step beyond programming languages, IEEE Computer 16, 8, August 1983, 57-69.
17. Terrenghi, L.: Design for Interaction in Instrumented Environments. To appear in Doctoral colloquium proceedings of Pervasive 2005, May 2005, Munich, Germany.
18. Weiser, M.: The computer for the 21st century. Scientific American, Vol. 265, September 1991.
19. Zhai, S., Morimoto, C. Ihde, S.: Manual and Gaze Input Cascades (MAGIC) Pointing. In Proc. CHI'99.