DESIGN FOR INTERACTION
IN INSTRUMENTED ENVIRONMENTS

Lucia Terrenghi

Abstract
Embedding technologies into everyday life generates new contexts of mixed-reality. My research focuses on interaction techniques supporting people who inhabit such augmented environments to continually make sense of the contexts within which they live, and to interact with virtual information embedded in the real world. To this end I work on the design of a novel interaction metaphor, creating a mapping between the affordances of physical real objects and the representation of digital information.

1. The Problem statement

Instrumented environments in ubiquitous computing [19] define spaces where technology is embedded so as to display and sense information in objects of everyday life. In this sense we have the chance to interact in a continuous display of information by moving in the real space and handling physical objects that are correlated to virtual information. Information migrates into the walls, where different appliances are invisibly interconnected. The lack of visibility and feedback bears the risk of a loss of control and awareness of interaction, and raises the need for new conceptual models.

My research aims to look at interaction techniques for instrumented environments, in the attempt to design and specify user interfaces that allow users to develop a consistent conceptual model, enabling them to interact with such an environment.

While interacting with the physical objects, humans develop a conceptual model relying on the explorative perception of the world and on the feedback received through all the senses. This enables us to understand how things work, and how they relate to each-other and to ourselves. Basic laws like gravity govern the physical space and allow us to build inferences of objects physical behaviors. But what laws govern the virtual environment of information and, most important, how users can make sense of it, is still a matter of research and design.

2. The Design Space

A first step in the design of user interfaces that allow for users' awareness and control, is the understanding of users' goals while interacting in instrumented environments. In this sense it is important to identify the main features that fundamentally distinguish the interaction in the space from the interaction with a desktop PC.

A first, obvious difference is that interaction is not constrained within a 2D visual interface, but rather distributed in a 3D continuum, which encompasses different stages in space and time. In this sense users' activities are much less confined and strictly definable, but rather evolve across multiple concurring tasks. As Lucy Suchmann [18] sustains, humans perform situated actions, i.e. in order to achieve their goals they perform certain tasks, according to the circumstances. Thus, situations determine their actions. To this respect the system will need to understand the users'

context, so as to adapt the output in such a way as to enable users to be aware of the context, and perform tasks that are necessary to achieve their goals in the given situation.

Additionally, the focus of human attention will likely be much more dynamic, shifting among several items/activities: even though desktop OS are mostly multitasking, i.e. allow several applications to run simultaneously, input focus has typically characterized them. This implies that a program needs to be selected to receive the next input from the user and there is only one active program at the time, while others can be open in the background. Given that human attention is limited, as well as our visual angle, input focus as we know it on the desktop will likely be revised or disappear in instrumented environments. Peripheral information thus becomes crucial in supporting users' awareness of the context. To support users' limited attention in an ecological way it will be necessary to go beyond graphical user interfaces and take advantage of the redundancy and consistency of multimodal interfaces.

In this set-up I foresee to develop a model encountering three main design aspects: information display, control and command mechanisms, and user conceptual model.

3. Hypothesis and Related Work

As in Winograd [20] there are three basic metaphors for how we interact with things, which therefore underlie the interaction metaphors we design: manipulation (hands, physical objects you do things with), navigation (feet, location and traveling), conversation (mouth, people you talk with). In order to cope with the goals above, I plan to verify the possibility of developing a conceptual model analog to direct manipulation, which suits the particular issues of instrumented environments. In such a vision I work on the hypothesis to develop an interaction paradigm that avoids the use of mouse, and relies on hand gestures for direct manipulation of information.

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 [17]. 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 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 remains consistent. 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 GUI of desktop metaphor provides affordances for mouse and keyboard input for interaction with virtual information, and maps to objects of the real world, that in turn provide affordances for gesture-based direct manipulation (*see Figure 1*).
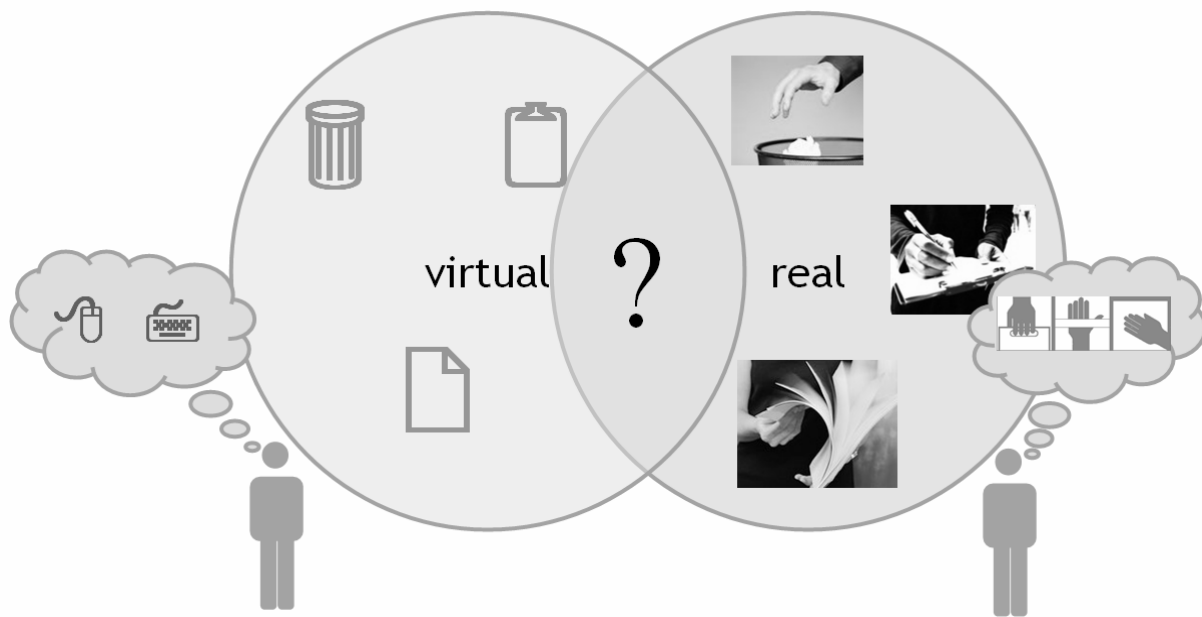
Figure 1. In the Personal Computer environment classical GUIs rely on metaphors in order to suggest the interaction operated with mouse and keyboard on virtual information. Real world objects provide affordances for manipulation. When mixing virtual and real information the issue of providing affordances for new interaction emerges.

My investigation focuses on the definition of general laws for interaction that can be applied in different domains and be supported by different devices, being aware that different domains (e.g. a museum, a living room, an office) can be populated by different artifacts.

This application-independent (in this case *appliance-independent*) metaphoric system aims at the definition of a general paradigm for interaction, that like the WIMP one allows the interaction with different domain-specific appliances, but better suits the ubiquitous computing settings.

With the emergence of ubiquitous computing scenarios, some work has been done in an analogue direction, looking at devices that can work as universal remote controllers [13] or support different functions depending on the way they are physically manipulated [14][3]. Other approaches look at natural input techniques such as gesture and gaze input [21] or to tangible user interfaces that work as tokens of information that can be manipulated in the physical space [8].

4. Main Design Challenges

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 provide feedback. Additionally, information may appear (or sound, or smell) 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. Objects' pliancy, i.e. their characteristic to be interactive, should be hinted [4].

Representation of virtual information can also be invisible, but it still needs to be somehow manifested in order for people to be aware of it, and control it. When lifting a bottle of opaque plastic we can recognize whether it contains some liquid by sensing its weight, without seeing the liquid inside it. If we shake it, we can recognize from the noise whether it actually contains liquid or sand. Even though visual representation might not be always necessary or appropriate, users need to be aware whether and where virtual information is present in the real environment.

Ubiquitous computing creates the possibility to augment the information related to an object that can be empowered with additional meaning and functionalities. Associating virtual information to the real world thus opens new possibilities for interacting both with virtual and real objects.

Objects of everyday life already carry information per se: their shape, color, texture, weight, temperature, material, all the aspects related to their physical features. Norman [11] applies the concept of affordances to every day life artifacts. 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. Ecological approaches [7] focus on perception and action based on human attributes: in this context affordances can be described as properties of the world defined with respect of people's interaction with it [5]. When seeing a glass we do not only know its function: we make an estimation of its weight, temperature, texture, noise and physical resistance even without touching or lifting it. On top of that, each individual builds a subjective perception, which relies on cultural settings and on the personal experience. In western cultures some people can distinguish a glass for Barolo wine from one for Chardonnay wine just looking at its shape, as this has a cultural semantic. In addition we can establish affective relationships to objects that relate to our personal experience in a certain symbolic way, thus differentiating the perception, and the association of information to an object: the same glass can represent just a functional tool for a user, or a gift for another one, thus differentiating the information and memories associated to the same object from different users. This requires for the understanding of the semantics of real world object and for playing with it in order to make a meaningful relationship between users and their contexts.

5. First solutions

Relying on the assumptions above, I am working on 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 play as interfaces, and hands as control devices. Ringel et al. [16] have worked in a similar direction, looking at direct hands-on manipulation without implements on SMARTBoards: differently from such approach, I am not trying 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 and gestures provide a much more varied one (e.g. press, draw a circle, point, rotate, grasp, wipe, etc.). Rekimoto [15] 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.

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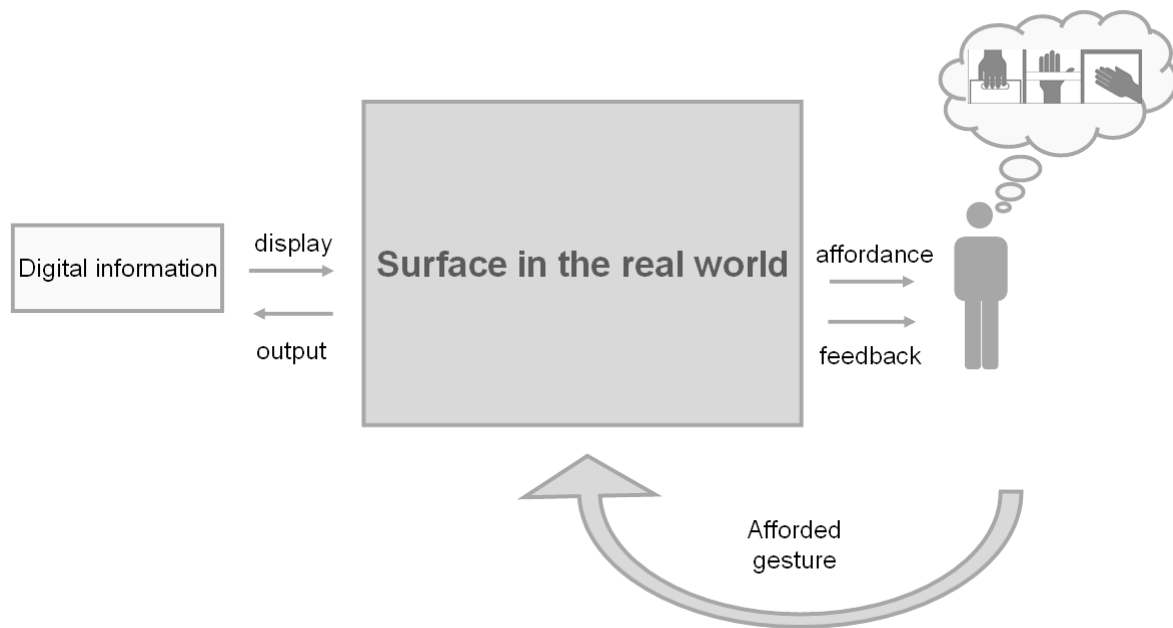
Figure 2. The representation of abstract digital information should present affordances for gesture-based manipulation when appearing on the surface. In such a paradigm, surfaces play as interfaces and hands as control tools.

A main aspect of affordances is that physical attributes of the thing to be acted upon are compatible with those of the actor [5]: in the setting 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 laid by rotating it with the dominant one. Representing digital information in an "affordable" way means to consider ergonomic aspects such as dominant hands, hands size, users' height, and so on. 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.

5.1 A metaphor for affordable interaction

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. Building on these assumptions I am working on the design of a metaphor that
- suits different environments
- is affordable for hands-based manipulation.

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. A first idea is 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 (*Figure 3a, 3b*). Empty mugs are expected to be lighter then 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 can be manipulated across the display. The non-dominant hand woks as command invocation, managing a menu of resources (e.g. drain, displays, printers): the dominant hand moves unites of information to the preferred resource (*see Figure 3c*). The pie menu appears in correspondence of a hand, 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 interaction.
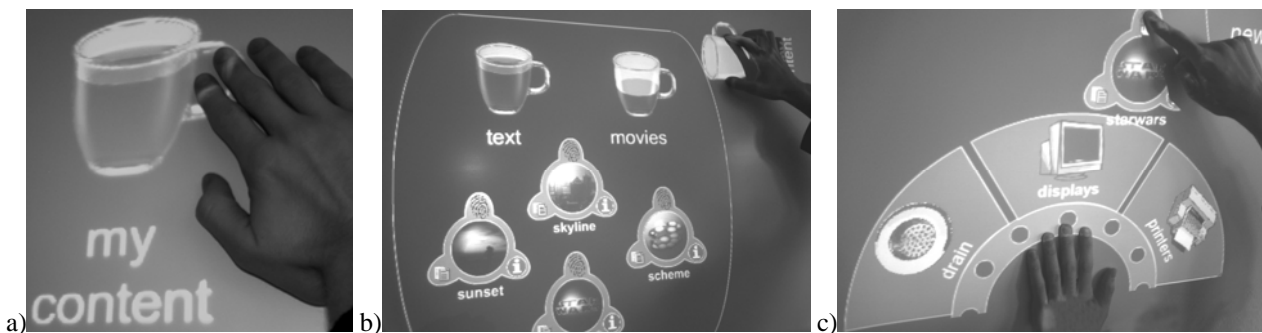

Figure 3. The mug metaphor interface. a) To move the mug/information container, the user touches its handle and drags it on the screen surface. b) To explore its content the user turns the mug. c) To cancel a unit of information, the user can drag it with the right hand to the drain displayed on the pie menu invocated with the left hand.

6. Process and Approach

My work develops in the context of the FLUIDUM project, www.fluidum.org, and can benefit of the infrastructure of an instrumented room, allowing for projection of graphical images on different surfaces, spatial audio display, camera based recognition of objects and gestures, sensors embedded in the surface of interaction. So far I've been conducting a deep analysis of related literature in order to specify the design space, identify the critical challenges and the different approaches, and the main issues that come in play when designing scenarios of use for such environments. While defining scenarios, existing work in the area of display-based activities, multiple displays management, ambient displays has been addressed.

A main point in the design of novel scenarios is to recognize users' goals and exploit the potential of novel technology. To enable people express their ideas and needs in such contexts, some first prototypes have been developed so as to provide a basis for discussion and interaction with users. The goal in this phase is twofold. On the one hand it is to identify *when*, *where* and *what information* people would like to access and be displayed in everyday life environments and activities. To this respect I have conducted a field study on display artifacts in domestic

environments. This work, consisting of 10 in depth interviews in 6 different households, has been based on contextual inquiry [1], cultural probes [6] and participatory design techniques of investigation in order to gather ideas. These preliminary results have generated some first design issues to be addressed when designing for ubiquitous computing of everyday life, and provided an insight on people's acceptance of ubiquitous computing scenarios in their daily lives.

In parallel, the design and prototyping of an interaction paradigm and user interface as presented in section 5.1 allows exploring requirements, both functional and non-functional (e.g. user requirements). In the FLUIDUM set-up, indeed, everyday life domains will be recreated so as to analyze how the virtual augmentation of such environments can support users' activities. In this sense existing appliances and prototypes are going to be integrated and interconnected so as to allow the performance of scenarios. Mostly, I am looking at scenarios of collaborative learning enhanced by a display continuous, and engaging users with haptic interaction.

Driving from scenarios of usage, thus providing an understanding of users' goals and domains, a set of most likely and common tasks can be extracted and selected, which will jet the basis for requirements definition. Interaction requirements will inform the design of the interface, will constitute a main source for the user interface formal specification, and will be used as assessment parameters for the evaluation phase.

Concerning the interface design, I aim to explore different modalities of information display, such as haptic and auditory displays, thus addressing non-visual affordances for the representation of digital information (e.g. the noise of dripping water for pending tasks). For the development of an interface specification I am going to select a formal model enabling to take into account all the agents of the interactions: to this respect I am looking at Interaction Frameworks [2] as potential tools for interdisciplinary integration of cognitive and system aspects of usability in design.

7. Expected Contribution

Much work has been done in order to make the system aware of the users' context by connecting sensors that measure users, environment, and domain parameters (e.g. body temperature, proximity, acceleration). Not so much has been done instead in the other direction, i.e. how to make the user aware of the system, and provide affordances for interaction in mixed reality. In this sense a main issue is to allow people to "sense" the space, so as to interact with it and recognize it as a place. Sensory motor theory [10] of perception has suggested some interesting work in this sense. The main account of this theory is that basically perception does not happen in the brain, seen as black box, but rather it is something humans do as explorative activity. For any stimulus, which can be perceived, there is a set of motor actions which will produce sensory changes regarding this stimulus. In TVSS (Tactile-visual sensory substitution) one human sense (tactile) is used to receive information normally received by another human sense (visual) [9]. This research promises to offer innovative ways to deliver awareness of the interactive context to the user, without affecting her focus of attention. To this respect, the design of multimodal affordance for the representation of digital information is a promising strategy for the achievement of users' conceptual model of ubiquitous computing scenarios.

8. References

[1]   BEYER. H., Holzblatt, K., Contextual Inquiry. Defining Customer-Centered Systems. Morgan Kaufmann, 1998.

[2]   BLANDFORD, A. , Barnard, P., Harrison M., Using Interaction Framework to guide the design of interactive systems. International Journal Human-Computer Studies 43(1): 101-130, 1995.

[3] 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.

[4] COOPER, A., Reimann, R., About Face 2.0: The Essentials of Interaction Design. Wiley, 17 March, 2003.

[5] GAVER, W., Technology Affordances. In Proc. ACM CHI 1991.

[6] GAVER, W., Dunne, T., Pacenti, E. 1999. Cultural Probes. Interactions, 6 (1), Jan. 1999, ACM Press.

[7] GIBSON, J. J., The Ecological Approach to Visual Perception. Hughton Mifflin, New York

[8] ISHII, H., Ullmer, B. Tangible Bits: towards seamless Interfaces between People, Bits, and Atoms. In Proc. CHI 1997.

[9] KACZMAREK, K. A.. Sensory augmentation and substitution. In J. D. Bronzino (Ed.), CRC handbook of biomedical engineering. Boca Raton, FL: CRC Press, 1995.

[10] 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.

[11] NORMAN, D.A., The psychology f Everyday Things. Basic Books, New York, 1998.

[12] PERRY, M., O'Hara K.: Display-Based Activity in the Workplace. In Proc. INTERACT'03.

[13] REKIMOTO, J. Pick and Drop: a Direct Manipulation Technique for Multiple Computer Environments. In Proc. UIST'97, (1997).

[14] REKIMOTO, J., Sciammarella, E.: TooolStone: Effective Use of the Physical Manipulation Vocabularies of Input Devices. ACM UIST Symposium on User Interface Software and Technology. p. 109-117, 2000.

[15] REKIMOTO, J. SmartSkin: an Infrastructure for Freehand Manipulation in Interactive Surfaces, ACM Press, CHI 2002, 113-120.

[16] 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.

[17] SHNEIDERMAN, B., Direct manipulation: A step beyond programming languages, IEEE Computer 16, 8, August 1983, 57-69.

[18] SUCHMANN, L.: Plans and Situated Actions. Cambridge: Cambridge University Press, 1987.K. 3.

[19] WEISER, M.: The computer for the 21st century. Scientific American, Vol. 265, September 1991.

[20] WINOGRAD, T., Flores, F.: Understanding Computers and Cognition. Reading, Mass.: Addison-Wesley, 1986.

[21] ZHAI, S., Morimoto, C. Ihde, S.: Manual and Gaze Input Cascades (MAGIC) Pointing. In Proc. CHI'99.