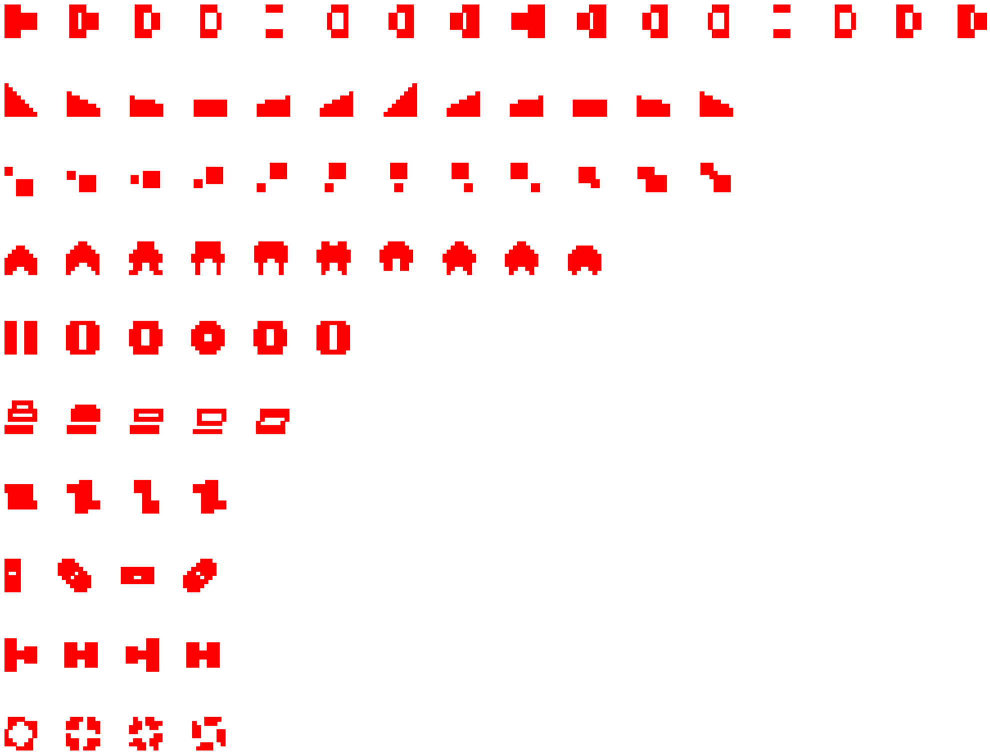
Interfacing Ambient Intelligence



Marius Hartmann

Department of Digital Aesthetics and Communication

IT University of Copenhagen

A thesis submitted for the degree of

Doctor of Philosophy

April 25 2005

Acknowledgements

I would like to thank the following people for their contribution.

My supervisor John Paulin Hansen, Associate Professor, Ph.D for his outstanding inspiration and support in making this thesis, and for conducting the statistical analysis reported in chapter 6.

Mikkel Holm Sorensen, Ph.D and Gert Balling, Ph.D for being such gifted colleagues.

Pieter Jan Stappers, Professor, Ph.D for giving me the chance to visit the IO Studiolab at TUDelft.

The DELCA team members.

My wife and family for their endless support.

Abstract

Mobile computing and positioning technology merge physical and virtual reality. This hybrid space requires us to look beyond conventional interface designs in search of new forms of interaction that will work in open-ended systems. In contrast to the safety of the virtual desktop, physical reality is open-ended. Challenges to the design of an interface that must function in a constantly changing world involve practical issues and defining new conceptual models of interaction. The practical, low resolution on visual displays with animation may be a viable solution to changing resolution, distortions, and scattered user attention. The conceptual approach taken is to avoid unnecessary encodings in the interface and instead build on causal relations between the interface and the immediate environment whenever possible.

The chief contribution of this thesis is the articulation of a visual format that is sufficiently robust, sufficiently swift, and sufficiently flexible to survive in dynamic environments. This approach is contrasted to traditional visual representations of agents and system functions.

Contents

1 Introduction **1**

1.1 Motivation . . . 1

1.2 Ubiquitous displays . . . 2

1.2.1 Physical-world display techniques . . . 5

1.3 Causal interaction . . . 7

1.4 Visualizing invisible Ghosts . . . 9

1.5 Outline of this thesis . . . 10

1.6 My original contribution . . . 11

1.7 How to read this thesis . . . 12

2 Context-dependent interaction **14**

2.1 Context as filter . . . 14

2.2 Context as a construct . . . 22

2.3 Guiding with mobile devices . . . 26

2.4 Summary of context-dependent interaction . . . 31

3 New interfaces **32**

3.0.1 Body, mind, and reality . . . 33

3.0.2 Software agents . . . 34

3.0.3 Augmented reality . . . 40

3.1 Sound interfaces . . . 41

3.1.1 Interactive sound art installations . . . 45

3.2 Interaction at hand . . . 47

3.3 Ambient displays . . . 49

3.3.1 Art as ambient displays . . . 54

3.3.2 Ambient displays on everyday objects 56
3.4 Ambient projectors 59
3.5 The push of information 64
3.6 Summary of new interfaces 68

4 Low resolution **69**
4.1 Grid formats 70
4.1.1 Technique 71
4.2 Nonhomogenous grid displays 74
4.3 Motion at low resolution 76
4.4 Mechanics and aesthetics of low resolution 82
4.5 Summary 85

5 Designing ambient agents **87**
5.1 The concept of DELCA 87
5.2 Ghosts 88
5.3 Visual format 90
5.4 The habitat of the Ghosts 98
5.5 Interaction with Ghosts 101
5.6 Causal relations to the Ghosts 103
5.6.1 Personal-display to public-display interaction 104
5.6.2 Public-display to public-display transitions 106
5.7 Ghosts on small displays 107
5.8 Summary of design proposal 110

6 Empirical evaluation **111**
6.1 The experimental setup 111
6.2 Experiments 1 and 2 reaction times and errors 112
6.3 Experiment 3 familiarity with voices 117
6.4 Experiment 4 figure-voice associations 117
6.5 Experiment 5 recall of figure names 119
6.6 Experiment 6 subjective evaluation 120
6.7 Experiment 7 blinking versus animation 121
6.8 General discussion of results 122

7 Toward nonencoded representation 125

8 Conclusions 135

A Animated Ghost figures 137

B CD-ROM 140

References 158

Chapter 1

Introduction

1.1 Motivation

This work investigates the design for integrating peripheral information across mobile interfaces and displays in the environment. I argue that the combination of digital information and physical location alters the prerequisites for interface design. One of the consequences of moving around in an environment is the change in actions one may perform. In the physical domain, our relation to environmental features afford possible actions (Gibson, 1979). In the conventional interface, we deduce the meaning of symbols from memory. The merging of these two very different ways of accessing information poses problems to people, as one mode requires distributed attention whereas the other requires focused attention. I look into the various ways this problem has been addressed by researchers and propose a new solution.

The majority of information people encounter when moving in the world comes from information displays in the environment - not from their personal device. Some personal devices that convey information related to the environment do exist, digital maps for example, but the use of these devices rarely augments the user's ability to interact with the surroundings. The mobile interface may isolate the user from the environment due to the need for focused attention or the modality used. Users of museum information systems, for example, are often quite unapproachable due to their use of headphones.

Is it possible to augment the information already available to us in our physical surroundings with virtual information without sacrificing contact with the world around us? In recent years, the ubiquity of digital information in physical environments has caught the interest of researchers across different fields. Ole Bouman of the Aarhus School of Architecture argues that the increasing potential for interaction between a given space and the people inside that space requires reflection as to the role of architecture itself (Bouman). Bouman sees the increasing spread of displays in physical spaces as a merging of digital information and architecture. According to Bouman, this "in-between physical and virtual space" forms a hybrid space that redefines architecture as an interactive medium. Bouman argues further that the visual possibilities of display techniques in physical environments may alter people's spatial perception of architectural structures. In this way, what constitutes space and location nowadays may come to be redefined. The dynamic situations created by this hybrid reform the functionality of architecture, requiring architecture to be adaptable and to provide dynamic situations. Bouman argues that the communicative aspect of architecture should be taken seriously into account with the use of display elements. Bouman also points to the new task of joining virtual information with physical space and how it may provide a novel interface (Bouman).

1.2 Ubiquitous displays

The spread of information technology into everyday life has put displays all around us. In supermarkets, for example, we encounter monitors, text tickers, and electronic price tags (Weiser, 1991). In this thesis, I deal with challenges posed to display systems in dynamic environments. The challenges may be divided into three main areas:

- Practicality challenges - how to supply the information.
- Interaction challenges - how to control the information presented to me.
- Sharing challenges - is the information meant for me or someone else?

The *practical* issues can range from cost to elements in the environment occluding the display. Of particular concern are liquid-crystal display (LCD) projections.

Even the physical size of a system must be taken into account. Back-projected displays may achieve high definition as large screens, but they take up relatively more space behind the screen. The temperature in the building and the power consumption of a system may also present practical problems. It is unlikely that one single technology will be capable of solving all the practical problems of a given dynamic display environment (Molyneaux and Kortuem, 2004). Instead, different display technologies may have to be combined in a hybrid display system.

Means for *interaction* are necessary in order to achieve dynamic displays. Interaction is vital not only for the individual session but also for the distribution of information to the relevant display. Naturally, the most desirable solution is one that maintains a direct relation between the display and the means for interaction. It is more practical to address displayed elements directly instead of manipulating secondary control panels.

In order to *share* displays with other people, a clear method for delegating displays is needed. It should be clear to people whether the information displayed is in response to their presence or actions, or whether they just happen to be viewing information intended for someone else. The sharing aspect is also crucial in order to be able to shift between displays according to the user's location. When the user moves, the system should be able to follow him by discreetly switching between displays.

Ubiquitous displays may change our interaction with computers, as many surfaces not originally designed for displays now may serve as interface for information. Mobile computing may enable information displays to respond automatically to changes in location. In this sense, people may wear an aura of information that affects their surroundings, and, vice versa, surroundings may react to the presence of individual users.

In this thesis, I argue that an augmentation of reality may be achieved by extending the conceptual model of the interface beyond screen estate. By crossing the physical limits of a screen device, visual and auditory information may be combined with displays in the physical environment, forming an *interactive space*. This interactive space is a hybrid space of virtual and physical information that may change the affordances of the environment. In the interactive space, a physical chair isn't just "sitable." It can be extended by the digital context into becoming part of a digital interface. When a museum guest sits in it, it could offer to tell him about the pictures in the hall.

Information displays in the environment are of a widely varying nature - from billboard-sized full-color displays to tiny monochrome displays on vending machines. One of the major points made throughout this thesis is that, if one wishes to use many display varieties in the environment, the visual design should be sufficiently robust against transformations by changing resolutions, distortions of perspective, and even semi-occluded displays 1.1.

Figure 1.1: The variety of shapes, forms, and sizes of displays we may encounter in the environment goes from tiny delicate screens to light bulb signs to projections on coarsely textured surfaces. From left to right, large low-resolution LED display on wall, wristwatch display, PDA, and projection on floor.

In this thesis, I propose a specific visual format for the appearance of such information. The specific design format was developed through DELCA (Disembodied Location-specific Conversational Agents), an IT University of Copenhagen project that intends to create concepts for ambient intelligent systems that people can understand (see the design chapter on page 87). However, I believe that changing resolution, distortion, and partial occlusion are challenges that any new display concept in interactive spaces should meet.

Resolution is an important issue in the design of ubiquitous displays. The visual format has to be sufficiently robust in order for the content to remain invariant at the various resolutions the different displays may have in an interactive space. Information could jump from one display to another, following the user as he walks. Some of the displays could be pointers he passes; others could be interactive signs; and sometimes, he might have to refer to his PDA to see the information. The network combining the environmental display systems may have a transfer rate less than that of a standard PC.[1]

1.2.1 Physical-world display techniques

Figure 1.2 represents an overview of physical-world display techniques. The overview introduces the concept of dynamic-mobile peripheral information, as viewed from the context of matrix attention-demand vs. physical form.

Using this figure, I address the failures of current mobile device interfaces and suggest what changes are needed to adapt them to interactive hybrid spaces. In their current state, mobile devices such as the mobile phone and the PDA rely heavily on the focused attention of the user, as per the well-known representational techniques of the WIMP interface.[2] However, the requirement of user

[1]A standard PC has a fairly high internal data transfer rate between applications and its graphics card (2.1 Gb per second with x8 AGP), but the same rate of transfer is unlikely in a local-area network. Molyneaux and Kortuem (Molyneaux and Kortuem, 2004) argue that high-resolution displays may cause their own problems because of the great amount of data that needs to be transmitted.

Molyneaux and Kortuem additionally propose data compression as a technical solution to the transfer problem. I believe this merely introduces new problems for data decompression and processing, as well as problems with acuity, as current compression standards involve lossy compression techniques (e.g., loss of data). This may result in the same visual blurring as previously experienced with anti-aliasing resulting from compression artifacts.

[2]WIMP refers to the composition of an interface by use of Windows Icons Menus Pointers. This type of interface was pioneered by Xerox PARC (van Dam, 1997) and adopted by Apple and Microsoft.

attention creates a gap between the device interface and the person's surroundings. If the design of a system requires people to direct their attention to the device interface, then the rest of the world is at least partly neglected. A person looking at the screen of a telephone is unlikely to discover subtle changes in the surroundings. If a change, however subtle, actually constituted important information meant for that person, he might find himself depending on the translation of the information into the conventions of the mobile interface to receive it at all.

The use of present-day mobile devices is often comparable to the use of a street map, which requires focused attention. The effect is a self-contained interface that lacks any direct reference to the world it inhabits. I propose a design concept that combines ubiquitous displays with mobile devices, which, in effect, makes them suitable for direct interaction with their surroundings. The proposal design changes their role from being static-mobile and requiring focused attention to being dynamic-mobile and requiring only awareness.1.2

ATTENTION

DEVICE	Focused	Peripheral
Dynamic fixed	Ticker, control panel	Landing lights
Dynamic mobile	GPS	Car alarm
Static fixed	Signs	Road signs
Static mobile	Street map, PDA, Mobile phone	Traffic cones

Figure 1.2: Overview of physical-world displays.

Because the dynamic-mobile peripheral modus serves as our goal state, I present more arguments for this modus and suggest how it may support a new type of interaction, "causal interaction."

1.3 Causal interaction

Perception of causality is a very important link between the user and information distributed to the environment. As demonstrated by Albert Michotte (Michotte, 1963), relations that are directly perceivable without previous encoded knowledge may consist of temporal timings only. If one wishes to produce an interface that does not require recognition of an encoded structure, this can be achieved by a causal display technique. An everyday example of a causal interaction that people have become used to is the car alarm. When people exit the supermarket, they tend to search for their car by remotely switching off the car alarm. With most cars, this results in both audible and visual confirmation that may be picked up within a moderate distance. The phenomenon of linking people to their cars via the car alarm is a representational relation based on causality rather than encoding.

> *Scenario: John finds his lost car.* Exiting the sliding doors, John notices that the parking activity has not yet reached a calm level. The lot is swarming with cars and customers. The air is filled with the noises of running engines, slamming doors, and the buzz of people. His mobile phone rings. John pauses to set one of his bags down and flips out his mobile. It's his wife. "No, honey.... Yes, I was just leaving...but.... Yes, see you there." While talking, he puts his hand in his pocket, finds the button for the car alarm, and presses it just once. Somewhere to the right, about 50 meters away, his car signals its presence: "beep beep." John turns his head toward the sound and confirms the location of his car by spotting the corner headlight just as it stops blinking.

This scenario illustrates how people learn to use technology beyond its original intended function simply because it is possible to do so. People make use of the tight temporal relation between the manipulation of the remote alarm controls and the car's display system to their own advantage1.3.

Finding a specific location by the means of a device would require some sort of mapping of the environment. The mapping would in turn need to be presented

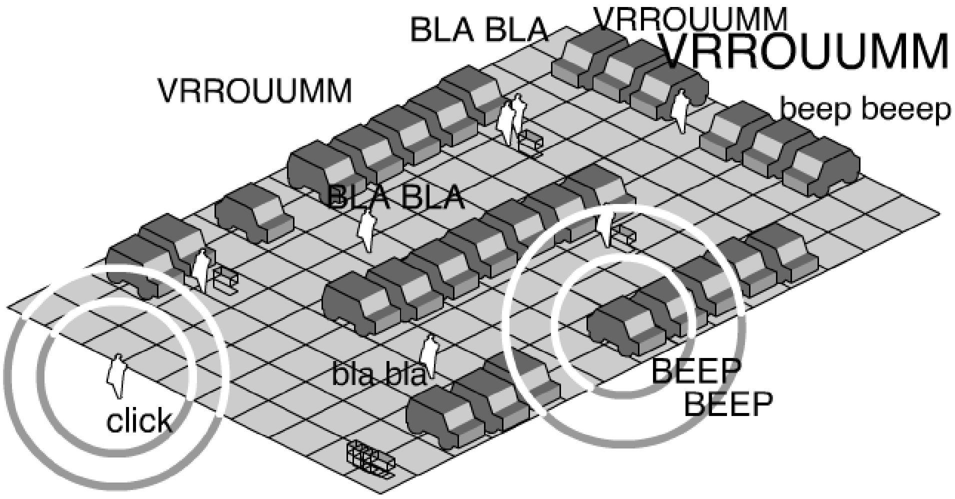


Figure 1.3: People may use the temporal dimension of interaction to organize a dynamic environment. In this case, a person uses his car alarm to locate his car in a crowded and noisy parking lot.

to the user by means of visual information. In order to interpret the mapping, the user would need bearings, landmarks, and perhaps even a representation of one's own current position on the map. The task could become quite complex with this approach, and the system would require awareness of location at a fairly precise level to point out elements in the surroundings. In particular, the depiction of the surroundings on the device would require some screen space and a possible demanding interpretation by the user.[1] Landmarks could ease interpretation of the map by referring to salient features of the environment. However, this would only be of value if the landmark database was kept up to date. In urban areas, landmarks easily become obsolete due to frequent changes in the physical environment: perhaps some scaffolding is temporarily occluding some otherwise-notable building, disguising its identity.

In the parking-lot scenario, this kind of uncertainty is handled with the simple push of a button. The rule John applies seems to be that whatever car immedi-

[1] The direction in which the device is pointing could also cause misalignment between the mapping and the surroundings, creating confusion for the user.

ately responds to the button is his car.[1]

I believe that the same basic rule of causality exercised in the parking lot should be applied to ubiquitous display interaction. Using signs in the periphery of attention on displays located in the environment may be a very effective strategy when coping with a dynamic environment. Displays located in the environment may address the user directly, without the need for referencing information.

I argue that the conventional way of dealing with displays often fails to overcome the limitations set by encoded representation (see page 125). The resulting interfaces come into conflict with the physical world because of their reliance on references to a learned meaning or value rather than to what may be readily experienced.

When multiple users share a public display system, causal display techniques may again reveal their limitations. Several car alarms going off at the same time, for example, certainly would create perceptual overload.

1.4 Visualizing invisible Ghosts

Several of the original designs presented in this thesis were problem-driven, specifically, the development of a coherent visualization of an ambient intelligence concept: DELCA (Paulin et al., 2004). DELCA stands for Disembodied Location-specific Conversational Agents. The DELCA agent concept combines context-specific conversationally mediated assistance with minimal and dynamic visualization, all designed to facilitate both engaging interaction and ambient calmness in combined virtual and physical environments.

The DELCA agent concept pioneers an interface in which services are assigned a personalized format of audio-visual characters. The main difference between the DELCA approach to that of traditional human-computer interaction (HCI) derives from its use of mobile location-based computing to form conversational, auditory interaction, supported by minimal visual cues.

[1] Car alarms are often a standardized beep, and as such do not contain referencing information that distinguishes them from others. This is important, as the identification of the car does not require uniqueness in relation to the car's visual and auditory signature but relies instead on the responsiveness of causal interaction.

The resemblance of the ghost characters to humans is mainly in the vocal domain; visual cues as to their presence are kept to a minimum, hence "Ghosts." The Ghost metaphor acts as an immediate explanation for the characters' lack of bodily presence in the visual domain, and it sets up a range of expectations for service. The ability of a Ghost service to follow the user around in a building, thereby seeming omnipresent, becomes easier to accept.

The combination of location detection and virtual characters allows us to engage services as we would people. If a person on his way from point A to point B is addressed by a service, he can choose to simply continue walking if he isn't interested in the service - it is a natural reaction to move away from things one doesn't like or want. But if the person does have an interest in the proffered service, stopping could initiate an engagement. In this scenario, the need for formal interaction by means of buttons and sliders is avoided and a causal interaction is made possible.

By nature, Ghosts are a bit unpredictable. They can sometimes disappear or become imprecise or slow in their service management. A basic user acceptance of these deficits will be of uttermost importance when ambient intelligence systems have to degrade on performance due to low bandwidth or imprecise tracking. Directional signs that come and go (see page 2) are also to be expected when they are provided by a Ghost.

1.5 Outline of this thesis

This thesis undertakes an investigation of the challenges to interface design posed by dynamic environments. Challenges described include elementary problems with the scalability of graphics, theoretical caveats, the technological problems concerning stability and bandwidth, and more.

I propose and initially test a visual design solution, as well as methods for interaction suited to hybrid space. In doing so, I present an introduction to the aesthetics of low resolution and establish a theoretical argument for non-encoded representation as a possible new paradigm for ambient displays.

This thesis does not cover the acoustic design of the Ghost figures, the system architecture of the location-positioning system, or its methods for tracking.

1.6 My original contribution

The major goals of this thesis include the following:

1. To identify new *design challenges* of ambient displays
2. To build the argument that a *causal-interaction* principle should be the cornerstone of ambient display systems that are context-dynamic and perceivable in the peripheral; then to apply this principle in my own design suggestions
3. To propose a coherent and aesthetic *visual format* for an ambient intelligence system. The format supports verbal, character-based interaction
4. An empirical study of user interaction with my interface design
5. To suggest sources of *aesthetic inspiration* for future combinations of mobile technology with architecture in everyday spaces

New design challenges The fundamental difference from dynamic environment interface design to conventional interface design is that the numerous possible relations of a dynamic environment defies the possibility of a predetermined mapping relying on encoding. The challenge to design is as such not only how to present a solution to this new situation, but also how to do so with respect to conventions of design that directly or indirectly relies on conventional symbolic references.

Causal interaction technique A way of handling relations without referencing to static values could be to utilize the properties of causality perception. This dynamic communicational technique allows for relations to be directly constructed and perceived without reliance of predetermined properties.

Visual format Low resolution is an economical and aesthetically viable design tradition. The coarse visual definition of low resolution makes it sufficiently robust to be visible across changes in dynamic environments. Low resolution may be further enriched by using the temporal dimension to increase spatial resolution.

Aesthetic principles The foundational aspect of communicating in dynamic spaces lies in understanding the world of human activities. The definition of the environment as an interactive space defined by the activities of its inhabitant has implications to the adaptiveness of aesthetic choices. More than position, size, shape or color, the timely display of information a fundamental quality of visual design.
My background lies with art and visual design. It is my intention that the figures and illustrations presented here should be given as much scrutiny as the text. It is my fond hope that their aesthetic and communicative qualities will be appreciated by the reader. This is particularly the case for the design of the Ghost figures.

In the list of figures, I have indicated which figures were designed by other sources. All other figures were designed by me.

1.7 How to read this thesis

Chapter 2 introduces the concept of context-specific interaction. I argue that context itself will become the filter needed most when information is everywhere. Context is not a fixed entity but constructed through the activities of humans. I explore one of the most promising areas of context-dependent interaction, physical guiding, by discussing previous prototypes for this type of activity.

Chapter 3 presents various designs that have been proposed for new interfaces for ambient technologies. Pros and cons of the different approaches are discussed.

In chapter 4 I reintroduce low resolution as an aesthetic visual format that has gained respect from outstanding visual artists in recent years, and which most importantly promises highly visible and scalable designs.

In chapter 5 the original contribution to the DELCA project is presented. This includes the design of animated Ghost figures and suggestions for how they should behave on various display platforms.

In chapter 6 the animated figures are presented to 40 subjects in an experiment to test their reactions. Some possible design improvements are identified on the basis of this experiment.

Chapter 7 is a conceptual reflection on the preceding discussions. It argues for a general movement away from encoded representation and toward immediate forms of representation that link interface and environment.

Chapter 8 contains the final conclusions.

Chapter 2

Context-dependent interaction

The last 50 years of research into Human Factors and Usability have identified certain basic principles of design - composition, contrast, and so on. These principles have subsequently introduced a number of conventions for how to structure the interface. This endeavor has been concerned primarily with the stationary media using text and symbols. However, the development moves toward distributing tasks to the surroundings, rather than keeping them within a centralized system. We already have seen a range of devices populating this new design space, such as mobile phones, PDAs, and augmented reality. New automated services, such as location-specific guidance, have been suggested. Projection displays, mobile displays, tracking, speech, and voice interaction allow for new forms of interaction, which in turn allows for new applications.

I argue that the concept of context may be used to filter out unwanted information and to avoid information overload. I also argue that filtering becomes especially important, as we cannot navigate within the high-resolution information space as well as we were able to with the PC.

2.1 Context as filter

Wireless devices and services are not new. However, navigating a physical space with ubiquitous digital services and displaying these accordingly produce new challenges for the portable contextual interface and its use.

In his well-known paper, "Some problems with the notion of context-aware computing," Thomas Erickson brings to light some of the complications and problems in using a computational approach for defining context (Erickson, 2002). He argues that, because of the multitude of possible combinations of context-related components - intention, location, interaction, and so on - a certain amount of common sense is required in any design algorithm. Yet common sense, he notes, is one of the major challenges posed to AI in general, and is therefore not applicable. However, by adding humans to the interaction cycle and letting the user react contextually, the computer's lack of ability to recognize context becomes insignificant. Erickson concludes that context awareness can be achieved by adopting a wider understanding of human-computer interaction: by allowing digital technology to gather and present information, and allowing the user to interpret and react to this information, the user becomes "part of the system" (Erickson, 2002) (p. 103). In this way, the strengths of both "technologies" - the computational and the human components - are exploited. In line with Erickson's idea of a user-system symbiosis in situ, I now expand on the interaction design of context-aware systems.

Location-based systems take the user's position into account when interacting with and presenting information to the user (Reitmayr and Schmalstieg, 2003). The fact that a user is situated in a train station may provide annotations to the interface for purchasing a train ticket. The interface can assume that the user wishes to depart from the station at which he is currently located. Our present location, in some situations, becomes an agent that deduces our intent 2.1.

The combination of mobile wireless computing with location detection offers contextual possibilities for extending information technology to the physical domain. This development spawns a number of challenges to the design of applications and the interfaces supporting them.

According to a study by Bertelsen and Nielsen, assigning interface values on the basis of physical surroundings can be achieved by applying bar codes to the various elements. When the user interacts with the elements, the bar code is read in and triggers an interface pertaining to the element (Bertelsen and Nielsen, 2000). In one example, a worker manipulates a motor through an interface on his PDA, which is connected to the main control system. The PDA interface is

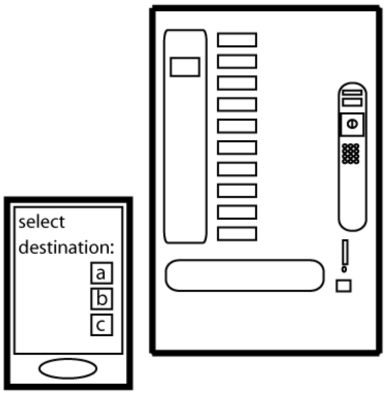


Figure 2.1: Devices in the environment. Features of the environment may influence the range of possible virtual actions. A nearby ticket machine may propose sales without the need for formal introductions.

triggered by a bar code on the motor. By ingenious use of standardized input mechanisms on the PDA, the scroll buttons can be used to control the motor. This enables the worker to manipulate the physical motor, while paying only minimal attention to the device interface itself, because the PDA is handled through unconscious, automated operations (scrolling up and down) (Bertelsen and Nielsen, 2000) (page 188) 2.2.

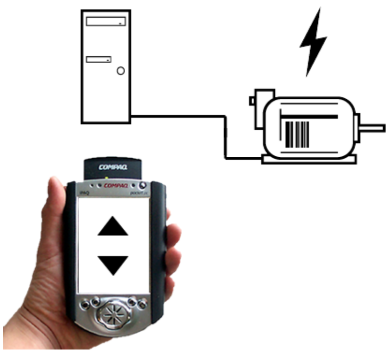


Figure 2.2: Bar code equipment. Worker operating engine euipped with bar codes by using scroll buttons on his PDA.

Augmenting the worker's action by crossing the visual borders of the interface to encompass functions and objects not represented on screen demonstrates a novel understanding of the interface. By relying on the context, Bertelsen and Nielsen establish a causal link between the device and its physical surroundings. The direct connection between interface and object of action eliminates the need for a representation of the action object on the interface. The interface manages the gap between virtual and physical reality by extending the interaction properties of the interface to that of physical elements. Location itself becomes both a physical and a virtual action that affects the range of actions the interface offers. In normal life, information is generally ordered by context and continuity. On the WIMP interface, it is ruled by representation.

The Semaphore project (Burrell et al., 2000) identifies two components that, when combined, unify invisible aspects of the environment to promote a contextual user experience:

- Context-aware computing
- Social navigation

Context-aware computing deals with the dimensions of location, user action, and the presence of other users to influence the information presented to the current user (Burrell et al., 2000). Interactions may be tracked by the system and associated with people, time, and location. Location may filter not only the display of information but also the actions that may be undertaken by the user.[1] In addition, social navigation makes information about previous user interactions available to the current user. This type of information allows the current user to make judgments about the environmental properties on the basis of other users' experiences. It may manifest itself qualitatively as user comments or a peer recommendation, or quantitatively as statistical information.

The Semaphore system likewise defines context as a product of three dimensions: location, time, and user. The system database contains entries with information about the author, location, and time that governed the circumstances at

[1] The approach of relating information to environmental factors is already found in the analog physical world as a way of distributing contextual information, for example, in tagging information to locations through the use of signs or as a discrete result of interactions in the resulting physical wear on paths or objects.

which the information was made available. The Semaphore system further supports social navigation by allowing for rule-based information entry and access. Information may be "shared" or "private." A level of hierarchy can be instilled by the system in combining location with that of data entry and access. For instance Burrell et al. describe how a professor could have designated areas in a classroom in which he could post information (Burrell et al., 2000) (p. 82). He could leave electronic notes around the lecture hall containing info on the coming lectures, or he could post grade results, available only to the relevant user 2.3.

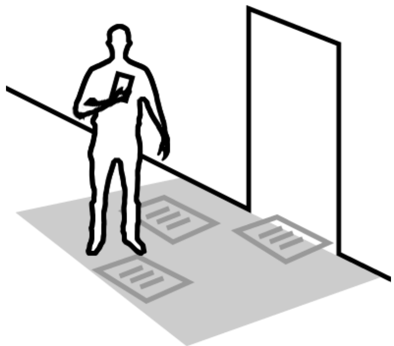


Figure 2.3: Public virtual notes. Conceptual illustration of notes left by a teacher and available to students outside the classroom.

Pascoe predicted that mobility will change the way interfaces are designed, and that mobility provides opportunities for promoting context awareness (Pascoe, 1997). Such a development cannot be exploited by transferring applications from desktop-interface design to the mobile device: it calls for a novel understanding of how to design applications that may intelligently interact with users on the basis of environmental awareness. Some of the contextual factors Pascoe mentions go further than just location and fellow-users' presence. Affective aspects such as blood pressure and body temperature of the user are also considered. Pascoe sees no limit to the kind of data that can be utilized by context-aware computing:

> In fact, any environmental factor that might influence the activities of the computer can be used, provided there is some mechanism for capturing it (p. 261).

Among the changes context may cause on interfaces and the related interface metaphors, Pascoe acknowledges "situated information spaces," as suggested by Fitzmaurice (Fitzmaurice, 1993). The concept of situated information spaces involves the mobile device as a "lens" revealing invisible information tagged to physical spaces and objects. The physical environment and the objects contained therein act as retrieval cues to the anchored information. This extension of the interface to conceive the physical environment as part of the interface is also recognized as a filter to avoid information overload on the limited interface of a mobile device. By utilizing the user's adaptation to the physical environment, orientation within the information space is facilitated. One can avoid cluttering information on the interface by distributing it to the wider physical horizon. In the "Stick-e Note" system (Pascoe, 1997), information remains invisible until the right variables are detected: a particular user profile, at a particular time, in a particular place. The Stick-e Note system works by distributing virtual notes around the physical environment. When placed, variables concerning "triggering" the note, such as proximity, time, and other measurable factors, are defined 2.4. Pascoe exemplifies the concept of triggering in a "go to the beach" note:

> if the location is the computing laboratory, it is a weekend and the temperature is above 25 (Pascoe, 1997)(p. 263).

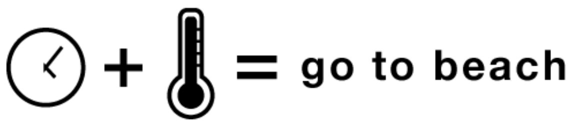


Figure 2.4: Triggering by context. System-detected variables trigger a message: if location equals "laboratory" and "temperature" is more than 25 Celsius, then "go to the beach."

The system computes context in an object-oriented system architecture. For instance, the context class within the system is structured hierarchically, thus allowing devices connected to the system to address information in a standardized manner.

Kruger et al. examined the differences between situations and the following variance of interface methodologies, from desktop to mobile computing, and propose a design called the BPN (BMW Personal Navigator) (Krüger et al., 2004). Their goal was to create a design that bears a similarity to navigational services - despite the different hardware requirements and contexts of the related three domains: desktop computers, in-car systems, and mobile devices (e.g., PDAs). They rejected the idea of a unifying interface for all three navigational situations because of the different level of complexity calls for "different user interfaces ergonomics." The concept of the design they proposed wasn't a unifying interface for the different technologies, but a connecting of the different technologies to achieve "a personal navigation service spanning different situations" (Krüger et al., 2004). The connection is thought to provide a seamless transition between situations in which the different technologies function, for example, the transformation from sitting before a desktop computer to walking with the PDA, then transitioning from a PDA to an in-car system, at all times maintaining unbroken consistency between interaction tool and actions 2.5.

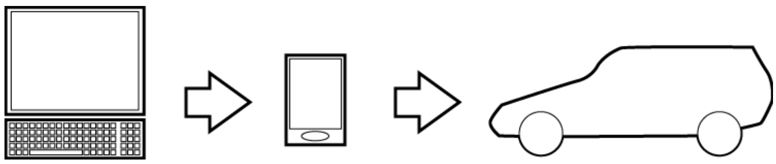


Figure 2.5: The BPN system of changing devices for different situations. From left to right, the desktop computer, the PDA, and the in-car system.

The BPN system by Kruger and colleagues links devices together by sharing data. It builds upon the concept of predetermining actions, i.e., that a degree of planning comes before actions. Working at a desktop computer is seen as a navigational preparation phase. The PDA and in-car systems carry out the preparation. This linear construction principle has its own drawbacks. Exploiting the different available modalities for the different situations seems a sound idea (e.g., using the higher-resolution display of the in-car system when available instead of the low-resolution PDA display). The use of the PDA as a consistency factor among the different devices, allowing for a continued interaction, also seems useful. However, the system runs a high risk of encountering the caveats of predetermined or computational context, as the system is built around the idea of carrying out a series of linear actions between various situations.

Efforts to extend the computational aspects of context-aware computing beyond that of location have been made (Schmidt et al., 1999) (Gellersen et al., 2000). Schmidt and colleagues augmented devices to sense such aspects of the environment as orientation and light. They also foresaw the potential for the device-utilizing sensor data to act as a precursor for whether to address the user or not. In this way, the context influences the information presented on the interface, thus reducing the problems of information overload. They argued further that a second wave of mobile computing had arrived with the introduction of "ultra-mobile computing" devices like the PDA, the mobile phone, and the wearable computer. Ultra-mobile computing is done while a person is moving and therefore should provide task-specific support.

Schmidt and colleagues produced a "working model for context" that maintains various aspects of context within a unifying structure. The model is divided into two main areas: human factors and physical environment. The human factors area is divided into the subareas of user, social environment, and task. The area of physical environment is similarly divided into the subareas conditions, infrastructure, and location. The physical-environment conditions themselves are also divided and subdivided: from "light" are derived level, flickering, wavelength, etc. (Schmidt et al., 1999)

In a multisensory architecture for context awareness, Schmidt and colleagues distinguished between two types of sensor, physical and logical. Physical sensors are hardware components characterized by their ability to measure physical aspects of the environment. Logical sensors are defined by their ability to gather sensory information within certain time cycles.

Picard and Healey investigated the role of emotions in human reasoning and how they may be incorporated by technology (Picard and Healey, 1997). Adding sensors that measure physiological data from the user's body that correspond to emotional states, such as skin conductivity and pulse, adds a dimension to the contextual information gathered by the device 2.6. The emotional information may function as a filter for what needs to be stored. The system may choose to record more information if the user dozes off, or less information if the user shows an emotionally strong response. The amount of information the device collects may thus be economized (Picard and Healey, 1997). The idea has yet to be tested in practice.

Measures of users' emotional states could be valuable in monitoring the information load dynamically. Most automobile drivers tend to lower the amount of distractions when faced with stressful situations; they tend to turn down the volume of the radio when engaged in heavy traffic, for example. Under such conditions of stress, a context- and emotion-sensitive system could filter away automatically any information that is not of immediate importance.

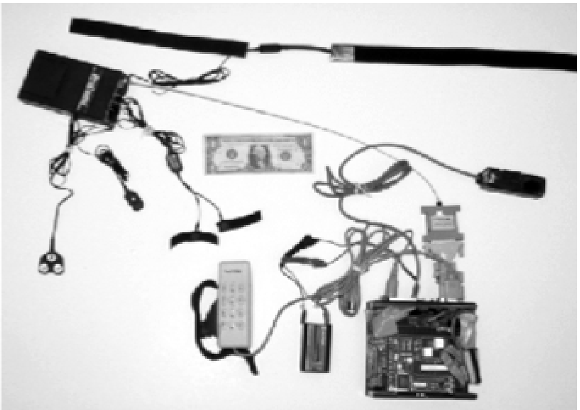


Figure 2.6: Affective wearable including a ProComp sensing system (upper-left corner) with sensors to measure respiration, galvanic skin response (GVR), and blood volume pressure (BVP), and an electromyogram (EMG) to measure muscle activity. Photo by Jennifer Healey.

2.2 Context as a construct

Perhaps a definition of context cannot encompass all situations (Dey, 2001). According to Dey, relevant aspects that generate context change with the situation at hand. Instead, he suggests this working definition of context:

> Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is

> considered relevant to the interaction between a user and an application, including the user and applications themselves (Dey, 2001).

Dey identified three aspects a context-aware device can support:

1. Presentation of information
2. Execution of a service
3. Tagging context to information

This categorization enabled Dey to formulate abstractions from which context-aware applications may be designed. Sensory data are gathered by widgets that in turn address the various sensory sources. Context interpreters within the system collect and transform information to suit the task at hand, for instance, by converting a name to an email address. The "situation" is deduced as an abstraction by a collection of states. The ongoing indication of states resolves issues concerning information-gathering, but it still leaves the designer with the challenge of defining the various triggers within the system.

Kurvinen and Oulasvirta argued that context is not a static property of the environment in which the mobile device exists (Kurvinen and Oulasvirta, 2004). Instead, they pointed to context as a dynamic construct that users continually generate. Kurvinen and Oulasvirta investigated context from a social-science perspective, and adopted the view that social context is what the user generates as he takes turns interacting. According to Kurvinen and Oulasvirta, previous research into context awareness has focused on quantifiable indictors of context such as the number of persons in a room, time, and location. They argued that these are static context factors, and that a causal relationship between data and context is presumed. As an alternative to raw data interpretation, they proposed the turn-taking approach.

The turn-taking approach deals with situational factors in context; the specifics of a given situation prohibit a general definition for context. Through their studies Kurvinen and Oulasvirta found that context is produced by user actions, and that these actions are in fact interaction turns. Context is never static but is dynamically constructed. The actions users undertake are grounded in a common understanding of relations within the group; users expect each other to act in certain ways according to the individual roles within the group. In addition, the multiple dimensions of context that can be acted upon sometimes can be incompatible with the current turn-taking.

Kurvinen and Oulasvirta also found that communication among users can develop group-specific practices in which communication may become uninformative to outsiders. However, the message may be rich to insiders because of the contextual meaning within the group.

Dourish pointed out that the problem underlying the difficulties in dealing with context in a computational setting originate from the representational nature of software systems (Dourish, 2004). Context cannot be reduced to specific variables that can be encoded and represented. Dourish instead offered an alternative model, one that takes user interactions and builds a contextual understanding of them.[1] Drawing on Suchman's theories of situated action (Suchman, 1987), Dourish (Dourish, 2001) emphasized that context is interactionally determined by the improvisational nature of user actions. A spatial or temporal triggering of contextual information falls short of producing context, in that the actions of the user don't necessarily follow a sequential organization that the system designer can predetermine or predict. Dourish's position suggested a connection between ubiquitous computing and embodiment: participation in the world implies a participatory status for all events occurring in the world - objects as well as actions. In this way, Dourish argued, the fundamental nature of things lie in their occurring participation rather than in their abstract value. The fundamental challenges of context-aware computing, he posited, can be uncovered through the investigation of the phenomenological understanding of embodiment. It is through the active participation in the world that things and events become meaningful. Focus of computational effort should change from *what sort* of contextual factors one may encode into the system to *in what way* the interrelations among these factors produce *meaning*.

Fitzmaurice investigated the linking of ubiquitous digital information with daily-life environments (Fitzmaurice, 1993). He used a small, portable display

[1]The idea of context as dynamically maintained through actions in an environment seems related to ecological psychology (Gibson, 1979) and "interactivism" (Bickhard and Richie, 1983), in that it rejects the encoding of dynamic entities.

connected to a location-detection device, which acted as a "bridge" between virtual information and physical objects. Physical objects are anchors, or hyperlinks, to digital information and allow for a spatial organization of the digital information in a real-world scenario. The display acts as an information lens that reveals the digital augmentation or structure related to physical objects, the "eye-in-hand" metaphor. Fitzmaurice's research indicated that location detection combined with a small, mobile display takes on the characteristics of a much larger static monitor:

- The movement of the display itself adds a temporal dimension of movement that enhances a 3D comprehension.
- The bodily sensation of the user's moving around extends the 3D experience.
- The responsiveness of the display to intuitive gestures the user makes creates a continuous interaction.

Bantre and colleagues explored the technical issues raised by allowing users to coauthor the situation context (Banâtre et al.). Mobile computing opens up the possibility for direct communication within a limited range between users. Users may create their own networks based on the proximity of other users 2.7. These networks are independent of a centralized WLAN system, and only exist temporarily on the basis of user movement. Bantre and colleagues' Spontaneous Information Systems (SIS) was founded on both the creation and the termination of encounters. The protocol for exchanging information among users has to be adaptable to the constant and sudden changes in the availability of information. SIS handles this by estimating the viability of the current link and dynamically selecting exchange rates and blocks of information to transfer, on the basis of the presumed "lifetime" of the connection. The same mechanism can be found in the typical reaction of people who experience unstable communication lines; people on holiday in far-away countries tend to focus on short messages, depending on the quality of the connection. They say, "we are fine," before elaborating on the weather. The SIS architecture doesn't require an infrastructure, as the system is dynamically composed of the user devices. Bantre and colleagues argued that the standard approach to ubiquitous computing contains fundamental constraints for natural interfaces, in that the architecture focuses on traditional human-computer interaction and user access to central information systems. They posited that the architectural demands for supporting this traditional approach are considerable, and that it doesn't take full advantage of the user device.

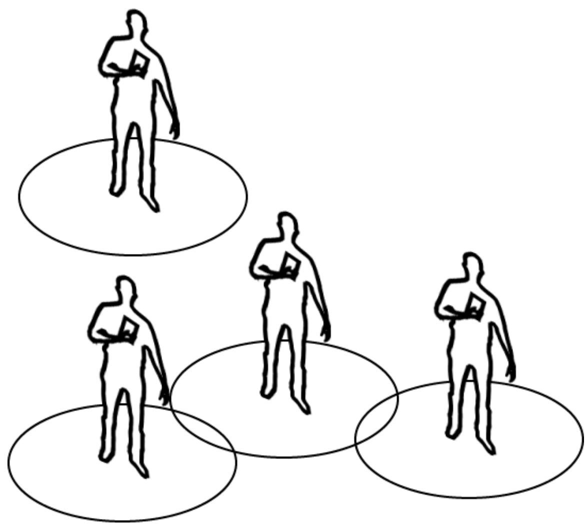


Figure 2.7: Ad hoc networks formed by users. Users form spontaneous networks based on the presence of other users.

The communicative assembly of multiple users in one location may give rise to virtual communities based on the physical and social interactions in the clustering. This clustering is directly rooted in the actions performed by users and defined dynamically.

2.3 Guiding with mobile devices

Allowing users to formulate the context of their surroundings is fundamental to creating a dynamic system. In the task of guiding people around a physical environment, being able to draw from the use of and social activities within that physical space is invaluable. The ability of a system to combine stable features of an environment with user perspective not only gives the details of the guiding procedure much higher resolution but also allows for a procedure that deals with the context of the tour. The route suggested by the system can deal better with the visitor's own purpose, as well as with the formal requirements of getting that person from point A to point B.

Guiding the user around a physical environment was also examined by Abowd and colleagues. They worked with tourists to create an example of users needing contextual information and directions (Abowd et al., 1997). Their approach to creating context awareness from location detection was mainly formed from working with location and the history of locations. The architecture of the Cyberguide system is modular, but this is not reflected in its interface:

> The system is an application that knows where the tourist is [and] what she is looking at, can predict and answer questions she might pose, and provide the ability to interact with other people and the local environment (Abowd et al., 1997)(p. 421).

Abowd and colleagues identified real-time wireless communication as a basis for the device. Wireless communication enables the device to interact with external systems and perform actions on the user's behalf, such as make reservations or buy tickets.

The Cyberguide system is divided into several components, each containing unique functions as a mobile tour guide. Abowd and colleagues assigned the system different roles to suit the particular guide scenario: the cartographer, the librarian, the navigator, and the messenger. The cartographer deals with map functions. The librarian deals with encyclopedia knowledge. The navigator takes care of positioning and orientation. The messenger controls communication. The modularity of the system enables the creation of hybrid services. For example, a historian may draw upon map, positioning, and library functions to create a unified picture. The Cyberguide interface does not use a personalized character for the different roles, however, but makes use of more conventional application design. Due to its pioneering state, the system was designed with off-the-shelf technology available in 1997, such as infra-red (IR) sensors and GPS.

The C-MAP approach to guiding combines an animated character with physical and semantic mapping to capture the user's context (Sumi et al., 1998). The C-MAP system is based on the Active Badge System (Want et al., 1992) for positioning the users and on PDAs with WLAN for communications. Sumi and colleagues also had tourists in mind as their target groups for the prototypes. The intention was to provide contextual guiding services to the users, as well as to provide communication features that may be used on- and offsite (e.g., from the internet).

The C-MAP visual interface is divided into three main components: physical mapping, semantic mapping, and interface agent. The physical map displays the geographical layout of the environment. The semantic map contains abstractions of exhibition contents. The interface agent "mediates interaction between the system and the user" with an embodied, animated appearance (Sumi et al., 1998). The agent is constantly visible, while the two other mappings are shown separately. The physical map is a 2D floor plan with markers for exhibition areas and the user's location. The semantic map contains "spring-loaded" boxes of text that change relations on the basis of user interaction. The relation between the semantic map and the physical map is maintained through textual links, from the semantic content to the exhibition area names. Recommendations from the system are communicated by adding extra "recommendation" icons to areas within the semantic and physical map, in relation to user interests.

The agent character can be one of 11 different user-selectable characters, each containing four states and several text messages. One character is visualized as a person with a teapot for a head (pouring tea from the side of the head). The purposes of the character are to signify the system's internal state, get attention, hurry up the visitor, and inform about and encourage use of the system. The agent resorts to four different actions: suggesting, thinking, hurrying, and idling. (The agent will display random messages concerning the general use of the system when idle.)

Sumi and colleagues learned from the user feedback that even though users experienced affection with the character, they did not experience improved system reliability. Sumi's group also acknowledged user feedback that indicated that the use of voice guidance together with visuals are effective means of communication.

It isn't obvious what added value the embodied agents brings to the Cyberguide system. The actions the agent may undertake don't seem to go well in hand with the intent of enhancing users' experience with context. Drawing attention to information contained in the maps by the agent seemed a bit odd, noted some of the participants. If the purpose is to direct the user's attention to map elements, why then redirect the attention by means of a central figure? Adding simple attention-getting elements would solve the design task.[1]

In the VisTA-walk system, the characters appear on displays in the surroundings, not only on the mobile display (Kadobayashi et al., 1998). The triggering mechanism for a character to jump between different displays is the system detection of the user's Active Badge. The VisTA-walk system's use of omniscient agents seems very promising and innovative; however, the automatic external displays that are based purely on ABS badge detection aren't necessarily relevant to the user, as the transformation of the character is based purely on predetermined factors (e.g., user location), and not on the actual interactions of the user.

One of the key elements in guiding is the ability to point to landmarks the user can perceive. Landmarks may become obsolete due to changes in the environment or occlusion. The constant update of landmarks' status may require overly vast resources. Yet if only stable features in the surroundings are used, such as architectural elements, the number of landmarks in the system may be too low.

Fagerberg, Persson, and colleagues developed a location-based system called GeoNotes (Persson et al., 2000) (Fagerberg et al., 2003), in which the users themselves are allowed to determine, define, and name "places." This bottom-up approach assures a high resolution of places relevant to the users and enables constant updates. Each mobile user can leave electronic notes at locations they visit and define new locations as a GeoNote. The information left behind is then

[1]Interrupting the user's activities and asking him to "hurry up" is not likely to be a feature that users will appreciate. Sumi's group explained that the need for this feature is due to the limited battery life of the device. This author would suggest providing extra batteries rather than pestering the user.

Additionally, random general messages put forward by the agent during idle periods are difficult to accept as a feature that would strengthen the contextual quality of the system. If the system isn't in use, changes are in favor of the system not being relevant to the user in that particular situation. Seen in this perspective, the four actions the animated character may display become redundant: "Suggesting" could be made within the framing of the map as the suggestion has relevance to the map. "Thinking" could be communicated conventionally through the "hourglass" type applied to the cursor, as this would be likely to catch the user's attention in case of need. "Hurrying" should be a feature set up by the user, not initiated by the system, and the profanity of such a feature could take many forms. "Idling" should be just that; in mobile computing, the system should remain ambient until relevant to user context.

readable by subsequent users. Additionally, it is possible to add comments to existing GeoNotes.

Users didn't restrict themselves to naming locations that were based on stable features in the environment. Some locations were named after events that had taken place there - one location was named after a child screaming nearby. One group of users even transformed the system into a Internet Relay Chat (IRC) system, naming locations after the content of the GeoNotes at those locations. Chat rooms were directly applicable to the physical environment. Physical navigation in the environment became a sorting mechanism for sharing and distributing information: by going to the cantina, users could then chat with other users in the cantina.

The challenges of combining the physical world with virtual mapping were explored by Benford and colleagues as part of their project Can You See Me Now? (Benford et al., 2003). Players navigating a mixed field of virtual representation and physical location played a mobile game of catch. The teams were divided into online and mobile players, with each group represented on the virtual model of the physical layout of the game area. The mobile players carried GPS for positioning. By defining proximity by a certain limit, players physically present in locations were able to "catch" the players virtually present.

One of the many interesting concepts in the "Can You See Me Now?" project is "adjacent reality": rather than try to achieve a precise correlation between the virtual and physical worlds, the concept allows for divergence between the two. Due to the physical nature of mobile gaming, it would call for enormous resources to maintain a well-defined model of even a modest-sized city. But because of technical problems and delays, even such a well-defined model would be bound to get out of sync with the users' real-world experiences at some point. Adjacent reality calls for player communication being what allows for the maintenance of context, rather than the computational system's construction of a concise environment. The different groups of players were given different and limited information, which is conducive to users' interpreting and exchanging the information. In doing so, the hidden divergence between the physical world and the virtual wasn't experienced by the users, thanks to the different modalities in use between the mobile and online groups of players 2.8.

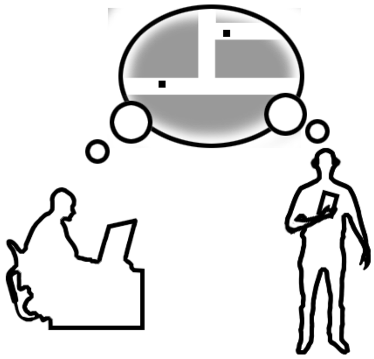


Figure 2.8: Can You See Me Now? The abstraction level of representation hides real-world versus virtual-world misalignment and allows "runner" and "player" to share a conceptual understanding.

2.4 Summary of context-dependent interaction

I have presented some previous work in context-dependent interaction. The experience from these projects has been that the context in itself may serve as a filter to reduce information overload. Context turned out to be a complex, multidimensional phenomenon. One promising approach for dealing with this complexity is to give users some freedom for constructing the context themselves. Other systems, such as GeoNotes (Persson et al., 2000) (Fagerberg et al., 2003) and "Can You See Me Now?" (Benford et al., 2003), have pointed out new ways of letting users co-author the context-dependent interaction. Guiding visitors has been a common scenario for several of the prototypes. Some of the systems, such as C-MAP (Sumi et al., 1998), ran into trouble trying to apply an agent metaphor for system services. Finally, the Vistatalk project (Kadobayashi et al., 1998) suggested how displays on mobile interfaces and in the environment may be combined in symphony.

In the next chapter, I shall explore how this combination can lead to a dynamic-mobile peripheral display.

Chapter 3

New interfaces

It may seem like a contradiction in terms to argue for both mobile and peripheral displays. Displays on a mobile device like a PDA or phone are so small that they require highly focused attention. In this chapter, I offer examples of how the limited interfaces of mobile devices may be radically extended. For this to happen, we may need to leave the 15-year-old Windows, Icon, Menus, Pointer (WIMP) format behind, and we definitely need to understand the specific requirements that unknown environments may put on our display design.

The evolution of the computer from mainframe to wearable device has changed the way we can use and integrate technology in everyday life. This change calls for new considerations of how to interface with and use technology. Wearable devices change relations to data from being hierarchically structured to being context-sensitive. This can have a tremendous effect on the design of system interfaces as Billinghurst and Starner pointed out:

> First, WIMP interfaces see interaction with the computer as the user's primary task. With wearable computers, interaction with the real world is the primary task. Second, WIMP interfaces assume the user has plenty of screen real estate and a pointing device. Wearable computers may have very limited screen space, and their input devices may have to be used with one hand, when out of the user's view, and at an arbitrary orientation - making pointing devices generally unsuitable. (Billinghurst and Starner, 1999) (p. 62)

The mobile user is no longer securely moored to the desk, and his hands may not be dedicated solely to a keyboard and mouse. Movement of the body hampers the precision of on-screen pointing, and interaction is often limited to the thumb or one free hand. The device's display is likely to be fairly limited and of considerably lower resolution than a monitor. I introduce some new candidates for mobile interaction.

3.0.1 Body, mind, and reality

The display must be readable at a (one-second) glance when looking at a mobile device (Marcus, 2004), as the interface becomes secondary to acting and moving in real-world situations. Most of the interface representations we have become accustomed to will not work in situations requiring highly distributed attention. For example, the use of simple navigational cues (arrows, etc.) may be more efficient than small detailed maps for navigating (Marcus, 2004).

Adding mobility to computing, the physical environment becomes a dimension of the interface, which must be accounted for. The location-based environment may offer both physical and digital services, as well as tasks to be integrated into the interaction cycle of the interface. Messenger applications can now offer information not only about *when* fellow users are present on the network, but also about *where* they are in the present physical sphere. Location detection proposes changes to the user's self-image and interaction space. The user's body becomes an important property of the interface in that physical location. This is a radical step that even the most advanced desktop interfaces have not taken yet.

In some computer games, the user must deal with a personal avatar when entering virtual worlds. The avatar mediates interactions with the virtual world, but the abilities of this "out-of-the-body" body are limited to the virtual physics of that world. The avatar may act like it has bodily experiences (e.g., vocal outbursts or indications of pain), but it has none and the user knows it. Interactions among avatars have grown very popular (e.g., in game worlds such as "Everquest"[1]), but the virtual world reduces our perceptual abilities, and it adds

[1] http://www.everquest.com

a layer of representations between our bodies and minds 3.1. With wearable devices, information technology is no longer a virtual space in and of itself, but is connected to our bodies and present in our physical space.

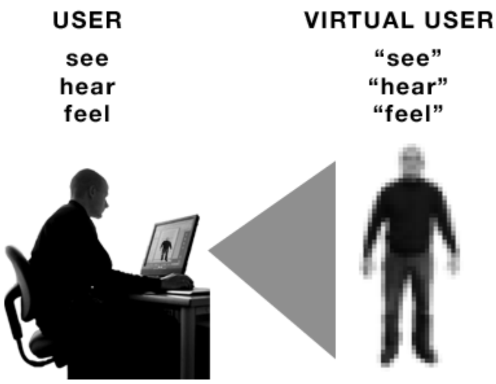


Figure 3.1: Interacting by avatar. The self-representation relaying the user's experience.

3.0.2 Software agents

Software agents are representations of system services in some human-like form. A well-known example is the Microsoft paper-clip, which gives advice to users apparently of its own will. The development of mobile computing has moved toward engaging or entertaining features that aim for a closer personalized relation to technology; this has spawned customizable features as well as characters that embody or reside in the products (Marcus and Chen, 2002).

Software agents that can filter and monitor events for the user could possibly prevent some of the information overload. SmartKom (Reithinger et al., 2003) is a uniform, multimodal dialogue interface that makes use of a "personalized interaction agent" called "Smartakus" 3.2. Smartakus is an anthropomorphic three-dimensional character whose visual appearance changes with the various screen sizes he adopts.

Even though these anthropomorphic creatures may be considered cute, it is doubtful whether presenting homunculi to deal with challenges to the interface is an optimal solution. The Mob-i device (Marcus and Chen, 2002) encompasses an

Figure 3.2: The Smartakus character. To the right, Smartakus can be seen giving directions via a PDA. Images from www.smartkom.org.

"affective personality" in the shape of an expressive animated character symbolizing the phone itself 3.3. The Mob-i creature is designed to resemble a mobile phone, and it resides on the display of a mobile phone. The idea is to cultivate attachment between people and their personal assistants or devices by adding a cute character residing in the device.

The combination of off-the-shelf technologies may result in new discoveries; the Mob-i's incorporation of a wireless earplug enables the design of services adapted to auditory messaging, such as text-to-speech and voice recordings (Marcus and Chen, 2002). The earplug doubles as a piece of jewelry and a source of information with a high degree of privacy, in addition to extending the Mob-i phone with wearable qualities - a hands-free interface.

Mob-i is capable of displaying a number of system states by facial expressions. The "reminder" message is a happy face, whereas the "low battery" indication is a tired, sad-looking face. Suppose you need to get a reminder message while the Mob-i is simultaneously low on battery power. What should the face look like then? The range of expressions a system would need to cover all possi-

Figure 3.3: The Mob-i personality. The various expressions displayed are, from left: normal, low battery, searching, alert, memory full, "thinking," successful/excited, and reminder. Image (Marcus and Chen, 2002). The design seems vulnerable to loss of detail at changing resolution.

ble state combinations seems too vast for a cartoon-like approach. In addition, ambient intelligence systems need to relate to external contextual states. Note how the low-battery indication lacks any degree of precision due to a dynamic symbol problem. If the symbol/personality can display affectations, should the current low-battery emotional state affect another of its current feelings, such as "successful/excited"?

One approach for creating an embodied interface agent was explored by Cassell and colleagues (Cassell et al., 1999), who argued that embodiment logically follows thinking of human-computer interaction in terms of conversation, because of the important nonverbal communication associated with conversation. The resulting character is "Rea," an anthropomorphic character that engages the user with conversation and gestures. The system is aimed at interpreting a number of user inputs such as facial expression, gestures, and nonspeech audio, using the same kinds of modality in expressing the computed output. This method is thought to extend the conversational interface by adding nonverbal cues and conventions as a social and psychological dimension of conversational interaction.

These "conversational functions" are aimed at strengthening the intuitiveness and effectiveness of the conversational interface. Cassell and colleagues saw the embodied character as being a necessary framework from which to present the user with gestures and facial expressions. The emphasis of the character's believability was placed on its ability to use human-like language rather than on its engaging visual appearance. Cassell's group acknowledged the possibility for a humanoid interface to raise users' expectations up to the level of conceiving of the computer as having human-like qualities, but proposed that this will only benefit human-computer interaction if the system supports a conversational interface.

In 1997 Ben Shneiderman and Pattie Maes warned against the use of anthropomorphized representations for digital assistance because they feared that it could mislead the user into believing that the agents possessed real intelligence (Shneiderman and Maes, 1997). The application of social traits to that of computers can be achieved with less than fully articulate visual agents. Nass and colleagues showed that users were willing to interact with computers as if they were distinct selves without any other assignment of human qualities than that of voice. Users were willing to treat computers as humans, even though they were aware this was not the case. Nass's group also discovered that the use of human voice alone is sufficient to induce this behavioral pattern and that different voices are treated by the users as distinct agents (Nass et al., 1994).

The sudden demand for honesty in user interfaces when dealing with software agents does strike as surprising when compared with conventional user interfaces. The desktop metaphor contains a number of tricks and betrayals that take place right in front of the user. Deleting a document by means of the trash can icon doesn't imply that the user believes in a real trash can hidden inside the machine. The trash can simply works as an easy way of setting a sector on the hard drive to zero (Tognazzini, 1993). Tognazzini described how a PC might be set up to utilize the relatively quick call connection of ISDN to simulate that the PC is always online, when in reality it only connects to the server as needed, thus saving on the phone bill. Tognazzini made a point about illusion and magic, in that we are ready to accept supernatural and powerful beings if we understand their limits. We are good at playing along and accepting outrageous characters, like magician entertainers, as long as we know the boundaries of the game. This has importance as to the choice of metaphor for the character and the honesty with which we introduce its capabilities. Tognazzini professed trust in the human ability to see through the agent:

> The computer is not capable of human intelligence and warmth; the character we create is. People will not end up feeling deceived and used when they discover, as they must ultimately, that the computer is nothing but a very fast idiot (Tognazzini, 1993).

Brennan and colleagues suggested that we simply acknowledge our daily-life interaction with disembodied beings:

> We should stop worrying about anthropomorphism and work on making systems capable of behaving as coherent interactive partners. Whether these partners are anthropomorphized or not, they should present their limitations frankly. People are used to dealing with many categories of partners: friends, strangers, the hard of hearing, disembodied voices on the telephone... (Brennan et al., 1992).

The point being made is that human communicational skills are socially adapted to adjust to the interaction partners' abilities, like children who adapt their own syntax when speaking to younger children. The same skill is what makes interaction with "dumb" technology like the computer possible in the first place (Brennan et al., 1992). The value of a personification of a service is dependent on a variety of factors, primarily, the context and related availability of modalities.

The enhancement or mix of physical and virtual reality by mobile computing has, in itself, implications to the choice of modalities for presenting personalized services. When one is situated in a physical space engaged with the environment, there is little need for a representation of one's own body. Neither is there a seemingly pressing logical reason for embodying digital services to such an extent that the question of their composition becomes an issue.

Like Brennan's and Nass's groups, I believe that users will be capable of distinguishing between artificial and human intelligence. The facial representation of agents applied to a mobile display has serious limitations because they require constant user attention and do not allow the user to focus on other activities. Moreover, an anthropomorphic visual agent struggles with the limitations of small-size displays, which makes it difficult to express subtle affective facial features. As an alternative, the DELCA approach uses the audio-language modality in timely combinations with discrete visual signs on the mobile device and in the physical surroundings.

The importance of facial expression in the application of synthetic faces to user interface was investigated by Judith Donath (Donath, 2001). Donath found that the timing and combinations of expression had to be very accurate so as not to cause confusion or unease. The awareness of context is especially necessary to interpret facial expressions (Donath, 2001) (p. 8). At the current state of technology, it is not possible to relay the nonverbal cues of human-human communication to a mobile device. This makes it impossible for the computer to perform social skills other than what may be deducted from interactions semantically.

Software agents have also spawned a fundamental discussion of the level of control available to the user. Patti Maes, a pioneer within software agents, has linked the introduction of agents to that of information technology by drawing attention to the special challenges posed by nonstructured information. She made the argument that the user's situation has moved from one of being in control over local files to one of being able to access networks with changing content. A combination of the nonstatic nature of networked information sources with that of a growing number of untrained users due to the ubiquity of devices will change the way we may interact with technology, from direct manipulation to agent assistance (Maes, 1994) (Shneiderman and Maes, 1997).

Ben Shneiderman and Maes insisted on direct manipulation as a viable solution to the growing amount of information. A display may work with thousands of entities signifying multiple relations - made possible, they argued, by the immediate feedback of the system and the "enormous bandwidth" to the user given by visual presentation (Shneiderman and Maes, 1997).

Both methods have enormous potential for a mobile interface: agents that may be used to delegate workload (e.g., to enable multitasking), and direct manipulation in the feedback loop and visual acuity (e.g., in dealing with large amounts of data).

Agent and direct manipulation are not to be treated as "either-or." In fact, the big issue is how users may interact with agents directly in mobile scenarios. Later I argue that our design must meet this goal to some extent (see p. 89).

3.0.3 Augmented reality

The head-up display has a number of advantages as compared with handheld devices like a PDA: weight, accessibility, privacy, nonintrusiveness, and the capacity for large images (Starner, 2003). Head-worn displays used for augmenting reality shows an image in front of the eye but allows for the physical background to be visible at the same time 3.4. As Starner explained, this causes an effect in which

> the computer's image seems to float in space, overlaid on the real world (Starner, 2003)(p. 15).

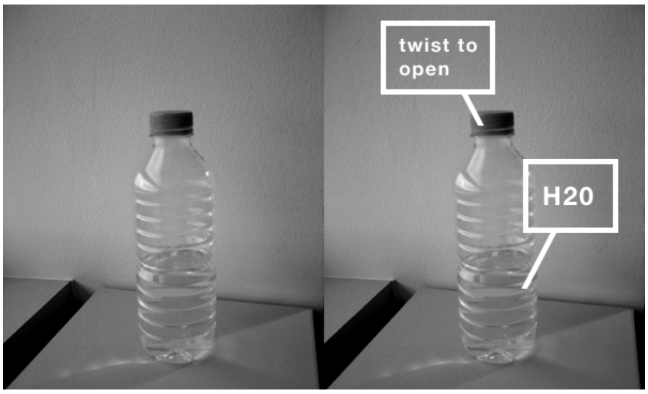


Figure 3.4: Head-worn displays reveal virtual information about the environment. To the left, normal view; to the right, augmented view.

The head-worn display may even provide a panoramic image or "virtual cylinder" when combined with sensors like a compass 3.5 (Billinghurst and Starner, 1999) (Starner, 2003). The virtual-cylinder concept is an innovative way of creating user interfaces that overcome the physical limitations of the window size by allowing the user visual orientation to control the display. This creates an intuitive interaction method for examining and navigating information, but the user is still present in the environment - he is not totally immersed in a virtual world.

The wearable device transforms information technology from being a virtual space in and of itself, to something connected to our bodies and a presence in physical space. Augmented reality enhances, rather than replaces, reality (Billinghurst and Starner, 1999).

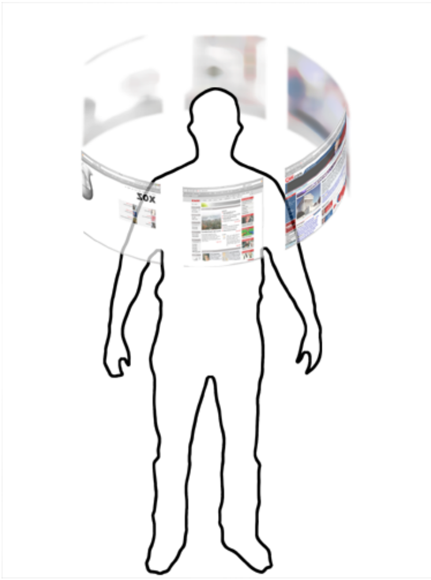


Figure 3.5: The "virtual-cylinder" interface experience. A user is centered within a cylindrical display of information through which the real world can still be seen.

3.1 Sound interfaces

Sound brings unique qualities to the interface. In vocal form it can serve as separate information layer,.for example, as commentary to an event, or it may support visual communication through abstract means, such as sounding an alert.

Sound is a very rich source of information and can induce strong emotions with users, as experienced in music. In fact, the effect of music can be so strong that it may change our very perception of a movie scene if accompanied by either scary or joyful music (Bussemakers and Haan, 2000).

Sound has populated the personal computer since the PC was first equipped with a speaker. We are all familiar with the "beep" of a booting computer and the greeting tune of an operating system. However, the use of sound in functional systems - such as office or administrative systems - has been modest in contrast to the use and importance of sound in computer games.

Sounds may imply meaning in much the same way an icon does. One type of sound used in interfaces is the abstract or musical called "earcon," for which real-life sounds are used as "auditory icons" (Brewster, 1998). An example of the use of sound in a daily-life interface is the traditional "dong" sound of a sports broadcast on television, notifying of a change in scores. Sound has the potential to inform the user about the availability of information without demanding full attention; in fact, the score information from the example above may be from a different game than the one being broadcast. The lower demand for focal attention has great value in situations where the user is already engaged in a task. For instance, the short audio notice of an incoming SMS allows us to maintain our engagement in our current activities and just mentally register the arrival of the new SMS.

In the Hubbub system (Isaacs et al., 2001) (Isaacs et al., 2002), sound is used to encode events and messages via an Instant Messenger system (IM). Users within a group are notified by sounds that other members have logged on, and users can even use sounds in messages to express their moods.

Hubbub allows users to maintain awareness of each other with visual and auditory cues. Each user has selected a distinct sound that is activated when their availability status changes. The "active" status has a distinct two-note sound, which precedes the unique user's sound. Users may be made aware that a particular user's status changed from "idle" to "active" by hearing the "active" sound followed by the "user identification" sound 3.6. The visual display contains more information about the user group, such as location and device used to connect to the system.

Isaacs' group suggested the notion of an interface that makes opportunistic interactions possible (Isaacs et al., 2001). They argued that interactions from daily life are based on meeting each other accidentally, which has important qualities for group behavior. The awareness of other group members' states isn't solely a functional matter from a transactional viewpoint. Awareness also has the important social aspect of feeling connected to the group. In situations in which team members are located at different geographical locations, the ability to maintain a certain level of common ground for the group is important. The constant sound update of other users' states in the Hubbub system enabled people to contact each other as soon as they noticed the opportunity for doing so.[1].

[1] Many contacts initiated just after login. A Hubbub user explained this behavior:

> [Sounds] are a trigger to tell someone something you've been meaning to ask

This spontaneous quality of the Hubbub system is also grounded in Isaac and colleagues' efforts to produce a "lightweight" interface that enables short response times in terms of both user interactions and system update.

One particular point of interest is the Hubbub notion of location; the feature of applying additional individual location information became a much-appreciated feature among users of the Hubbub system (Isaacs et al., 2002). Users not only applied location information but also used the interface space to post amusing or contextual information - what the weather was like at their location or special information in very specific detail, for example, snow depth at the resort.

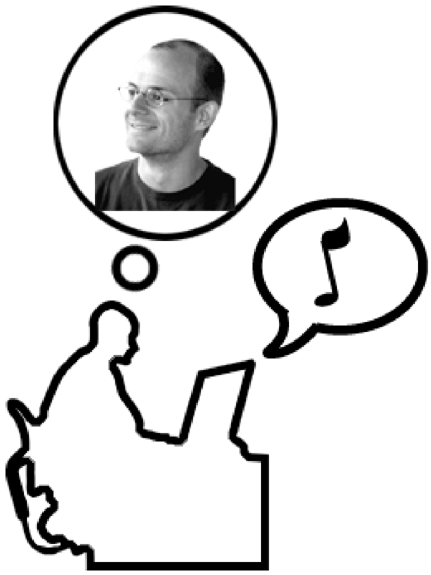


Figure 3.6: Using sound as a notifying cue. Sound causes the user to become aware that "College Joe" is online.

Brewster proposed the use of nonspeech sound in voice interfaces to convey properties of menu structures in telephone-based interfaces (TBIs) (Brewster, 1998). A set of sounds was tested on users to identify the ones most easily remembered. The ones most frequently recalled were the used to study an application of nonspeech sound representing menu hierarchy information. Brewster

(Isaacs et al., 2001)(p. 183).

found that earcons provide good navigational cues in TBIs, extending the use possibilities of auditory communication 3.7.

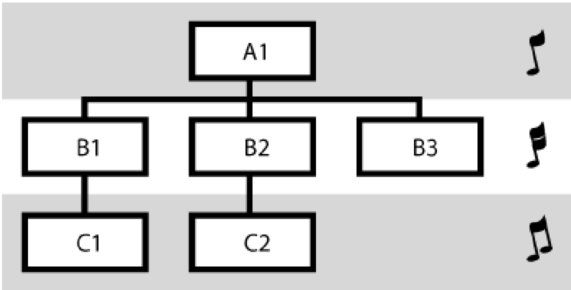


Figure 3.7: Communicating structural location with sound. Sounds may communicate location within a structure.

Brewster argued that the use of sound may solve some of the problems related to the small display size of mobile devices (Brewster, 1998). He tested the use of sound in combination with size of screen buttons on PDAs, and found that button size could be reduced significantly if accompanied by sound. Brewster also found that the test environment influenced the results. The conclusions drawn from laboratory testing didn't match those from testing in real environments. The dimension of mobility in a real usage scenario had a considerable effect on interface usability.

Brewster, like Billinghurst and Starner, posited that most design problems on the mobile interface stem from the Windows heritage, and in particular he emphasized the use of sound combined with small interfaces, so as to create an interface that doesn't require the user's visual attention (Brewster et al., 2003).

In addition to these experiments, Brewster and colleagues created a 3D sound interface (Brewster et al., 2003). Subjects equipped with stereo headphones and position trackers were able to use head gestures (i.e., nodding or shaking) to interact with an invisible interface laid out before them in sound 3.8.

"The 3D sound interface" is an audio version of the virtual cylinder, but spread out in space (Starner, 2003). Brewster and colleagues' idea of using sound to interface to invisible structures is an interesting solution to the problem of mobility and visual awareness. 3D sound does not require the user to pay visual attention to the interface while moving around in the physical environment. Thus

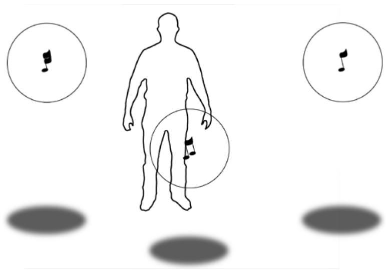


Figure 3.8: Using 3D sound to create audible structures. The user may experience structures in space based on a spatial perception of the originating sound.

the system is called "Eyes Free" (Brewster et al., 2003). Three-dimensional sound and design form a "soundscape" of auditory icons (e.g., the sound of rain, themed tunes, car horns) from which the user navigates.

3.1.1 Interactive sound art installations

The sound interfaces described above may have immediate use in the home and in organizations. For inspiration, I also present recent, purely aesthetic works with sound interaction.

Combining sound with visualization in the physical environment requires careful timing. Levin[1] and Lieberman investigated the aesthetic possibilities for making the human voice visible by so-called in-situ speech visualization (Levin and Lieberman, 2004) (p. 7). Augmented-reality techniques combine sound and visualization, coincident with a presumed origin. The prototype, "Hidden Worlds," was developed with projection displays, 3D goggles, and computer-vision-based tracking. In the installation, a group of users were equipped with tracked 3D goggles. The users were then able to experience others' vocal utterances as visual abstract entities deriving from the speakers' mouths 3.9. These "sound gestures" are visually defined by the volume, duration, pitch, and timbre of the individual user, and float from the mouth into an orbital pattern around the users' heads in a flock-like fashion. Their design draws from the "speech bubble" tradition used in comics, and was further inspired by early phonetic symbolism as developed in Gestalt theory.

[1] Golan Levin is a former student of John Maeda

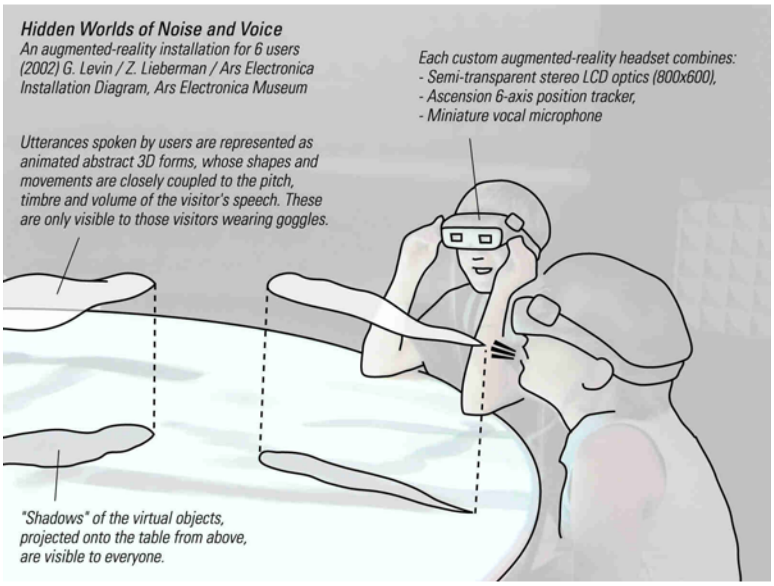


Figure 3.9: Hidden Worlds users wearing goggles and "seeing" each others' sounds. Only the shadows on the table are perceivable to outsiders. Image (Levin and Lieberman, 2004).

The RE:MARK system developed by Levin and Lieberman featured users speaking into a microphone (Levin and Lieberman, 2004). The system recognizes phonemes and writes them to a screen in front of the user. Phonemes not recognized are displayed as abstract shapes. The shapes become more pointy and irregular for high-frequency spectral centroids. The visualizations are projected from a shadow of the user, and the shapes may be pushed around by moving the shadow.

Findings from the RE:MARK system were expanded upon in the "Messa di Voce" system, which provided the same techniques for full-body interactions. Visualizations were now susceptible to the pitch in the user's voice, and the visual effects were thereby controllable. In a performance by an artist Levin and Lieberman further added the capability to interact with the visualization and replay the represented sounds. A shape is formed by a sound; touching the shape will replay that sound 3.10.

Research by Levin and colleagues (Levin and Yarin, 1999) (Levin and Lieberman, 2004) in the intuitive correspondences between sound and images may be

Figure 3.10: Messa di Voce installation. The user's voice is visualized by graphics on a large screen. Interacting with the graphics replays the related sound. Image Levin and Lieberman.

of great practical value. Images are generated by sound, and sound is then regenerated by the manipulation of those images. Even though the user maintains the sound-image relations, they are not based on encodings. The references between visual and audio representation are kept to the spatio-temporal relations between users and images. The relations are defined in the causal process of users uttering sound and seeing it visualized.

3.2 Interaction at hand

Levin and Yarin presented the acceleration-based interface as an alternative to traditional key and pointer solutions (Levin and Yarin, 1999). The acceleration-based interface lets users manipulate the device by movements. The changes in direction (e.g., tilting of the physical device) are then detected by an embedded accelerometer and transformed into manipulations of the interface. According to Levin and Yarin, this type of interface can produce expressive nuances, as well as provide continuous input. They investigated how to make drawings on a small display, but the principle can be used for other tasks as well. In a prototype, Levin and Yarin acknowledged the aesthetic potential of an 8x8 display; however, their goal directed them toward higher-resolution displays 3.11.

Intuitive actions in the physical world that influence the interface directly (e.g., by shaking, moving, or tilting it) enable direct manipulation of the interface without the need for buttons, sliders, or selectors mediating the actions

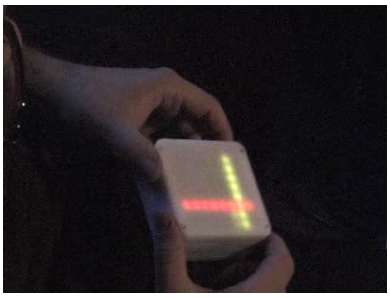


Figure 3.11: Levin and Yarin's. prototype of the acceleration-based interface. The lines move along the x and y axis when the device is tilted. From "Bringing sketching tools to keychain computers with an acceleration-based interface (Levin and Yarin, 1999).

complicating the interaction.

Ishii and Ullmer constructed the concept "Tangible Bits" for how to use physical objects as an interface (Ishii and Ullmer, 1997). They defined Tangible Bits as basically having two states or functions:

- Allow users to grasp and manipulate bits (digital information) with their hands at the center of attention, and
- Enable users to be aware of background bits at the periphery of perception using ambient media.

Tangible Bits form an interesting part of a direct-display system, in that the interaction required to interface information is mapped directly to objects. Opening a bottle is a direct interface to accessing the its contents; this is due to the visual, auditory, and tactile experiences of removing the cork3.12. Tangibility encompasses the extension of the interface beyond the screen with virtual actions accessible through environmental features.

Ka-Ping Yee further investigated the principle of moving a tracked display around, manipulating a virtual workspace (Yee, 2003). Using his "Peephole Display," Yee extended the "Chameleon" idea by combining it with PDA techniques for manipulation, such as pen interaction, and other principles, such as zooming. Yee recognized that his early prototype approach of limiting the interface to that of 2D brought it closer to working order, in much the same way as the "Toolglass"

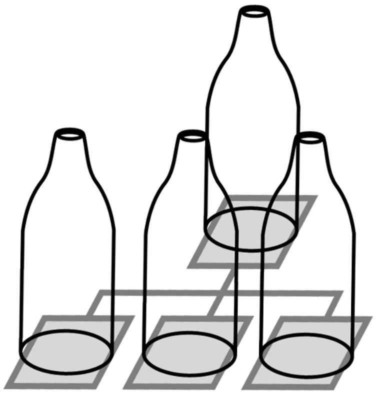


Figure 3.12: Tangible Bits mapped to digital information, with the objects themselves becoming interfaces. The environment as a whole enables focused attention to the graspable media, as well as awareness in the shape of ambient displays.

approach (Bier et al., 1994), in that it doesn't contain a "real" 3D mapping, but creates the illusion of a fixed virtual landscape by scrolling the virtual information counter to the PDA's physical movement. By incorporating location awareness, the Peephole Display creates the same direct mapping between virtual information and real-world position, and, despite the mapping being only 2D, enables the qualities of movement and spatial memory.

3.3 Ambient displays

Integrating devices with the environment in a nonintrusive yet easily perceivable way is a fundamental challenge to ubiquitous displays. This fundamental challenge comprises three distinct components:

- Displays should be adaptable and robust; that is, information should be communicated seamlessly through several displays in the environment, and interactions should be maintainable after interruptions.
- The displays should enable people to change readily from being aware of the information displayed to interacting with it.

- The displays should be customizable to the individual user and thus be capable of alternating between the different information sources (Elliot and Greenberg, 2004).

In their pioneering work on ambient displays, Wisneski and colleagues envisioned architectural space as an interface to digital information (Wisneski et al., 1998). They concentrated on how to present subtle changes in the environment that could convey digital information. They insisted on subtlety in expression, which inspired them to create an almost poetic set of ambient displays. They worked with the type of sensations one might encounter in the outdoor environment - rain, wind, and temperature - which do not preclude other activities. Wisneski's team saw the Window metaphor as a confining interface to information. In the WIMP interface, the user can only address information through a narrow view and only by keyboard and mouse. This method infringes upon the user in two ways:

- The presentation of content is mainly visual, with the visual presentation itself subject to limited screen real estate. Communication is encoded (e.g., symbol-based), which means that the experience (source of information) isn't readily perceivable but must be interpreted.

- The interaction available to the user is limited by the physical tools: the mouse and keyboard. The interactions require substantial visual feedback, and rely on tangible qualities to only a small extent.

But how can we perceive multiple information sources without the risk of information overload? Wisneski and colleagues argued that the physical environment itself enables a perceptual richness emerging from the variety of possible modalities - form, movement, sound, color, smell, temperature, light, and so on. The diversity in the possible channels for expression enables novel interface techniques that may effectively display multiple information sources, whereas the conventional Window interface suffers from lack of screen real estate, which causes the various information sources to compete for attention 3.13. The escape from the confinement of the screen becomes even more compelling when everyday interaction with information technology must take place while other off-screen

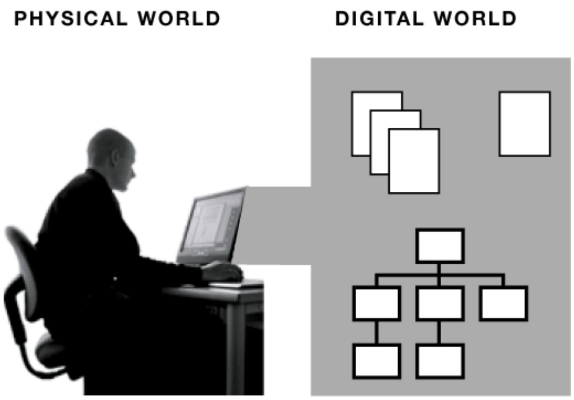


Figure 3.13: Schematic showing the conceptual and perceptual confinement of Window-based interaction with information, as argued by Wisneski and colleagues. The perceptual modality is purely visual, communicating abstract representations of information hierarchies.

tasks are performed (such as human-to-human interaction). Wisneski's group acknowledged the need for constant monitoring of information for specialized tasks in specific environments, such as cockpits or power plants, and they argued that these are examples of existing ambient media that have been developed to present constant monitoring of multiple streams of information.

The "Ambient Room" by Wisneski and colleagues was created from the radical idea of putting the user into the computer, to create a controlled physical environment a person could inhabit, while still maintaining relations with other people. The room itself would be fitted with numerous displays controlled by different modalities. One display consisted of water ripples reflected onto the ceiling. The ripples were produced by a small solenoid mapped to a hamster wheel outside the room, which relayed the hamster activity level to the Ambient Room. Knowing the activity level of the hamster is not essential, and the example could easily be used with other information sources, such as the activity level of an infant sleeping in a room next door. Another display featured the sound of dry-erase pens whenever the meeting room whiteboard was in use, signifying meeting activity to people outside.

Getting the signals from the displays to balance well with the environment turned out to be a major design issue. Initially, a vibrating device represented the hamster, but when the vibration technique proved annoying to users, it was replaced by the water ripple effect. Sounds of natural events turned out to work best, as users tended to get tired of unnatural sounds.

Active perception of our environment affords us the ability to focus our attention on its different features. Wisneski's team argued that the subtle displays should be able to coexist with the other tasks that take the most of our attention. Our ability to switch between perceived events enables us to quickly bring displays that are present in the background of our awareness to the fore of our attention. The design of the environment should support an effective, seamless transition from foreground to background.

Wisneski's group identified a threshold between foreground and background awareness based on the level of change of the ambient display in relation to the usage situation. A display may be well balanced when used singularly, but the same display may be too intrusive when used in numbers.[1]. "Calming" information down too much could jeopardize the perception of important events. Wisneski and colleagues observed that people adapt to given environments and learn to perform with increasing efficiency. Perhaps the information overloads connected with ambient noise level could be managed the same way we learn to acquaint ourselves with the creaks and strange noises of an old house.

According to Wisneski and colleagues (Wisneski et al., 1998), the basic idea of networked devices in our environment displaying a virtual layer of information is a most promising one. In this way, surficial affordances may be enriched by ambient information. If care is taken to integrate this concept into our surroundings, we may learn to distinguish between an air conditioner that can provide a weather forecast and an ordinary air conditioner.

Wisneski's group stressed the importance of selecting the right modality for various situations. A person engaged in visual tasks should be addressed by auditory means, and vice versa. However they also pointed to the problems of switching dynamically between different modalities during the course of use. Users may be confused about information mapping when switching from one

[1]This can be experienced in a shop selling clocks - at noon, the otherwise calm time displays suddenly conspire to extreme intrusiveness, whereas when any one clock is on its own, the noon signal is more quietly realized.

modality to another, when sharing displays with other people, or when different information sources are presented on the same display. This is a substantial design problem that I address in the chapter, "Designing ambient agents" (see p. 89) concerning the development of causal interaction within the DELCA displays.

Displays scattered around in the environment present a design challenge of how to present information seamlessly among multiple monitors (flat-panel displays, projections, etc.). Hutchings and colleagues reasoned that the increasing affordability of displays challenges the notion of computing based on a single-display stationary device (Hutchings et al., 2004). They suggested that this change implicates the development of new interaction techniques for multiple-monitor systems as well as new applications for environments rich in displays. Hutchings' group also examined displays in the environment used as "peripheral information-awareness aids" (Hutchings et al., 2004) (Stasko et al., 2004).

> An umbrella in a scene might represent traffic on a local road and be colored green when traffic is moving well, yellow when moving moderately, and red when moving slowly (Hutchings et al., 2004) (p. 5).

The work of Hutchings and colleagues was bounded to a personal workspace defined around a stationary location. They did not address issues concerning displays that change according to the user's location.

Stasko and colleagues' "Infocanvas" project (Stasko et al., 2004) examines the use of a display framed like a painting to convey information encoded in artwork figures. For instance, they suggested variables in a painting, such as the color of the sky, the alignment of objects within the composition, and so on, to be controlled by the property of an information source linked to that variable. The height at which a bird is positioned within the Infocanvas may be defined by the current state of the stock market.

The Infocanvas seems more like a screensaver than a new tool for communication; perhaps this is due to the use of nave figurative art. Changing the position of a bird in a painting may refer to a changing value, but only in a single dimension. When mapping the stock market, one would first need to know the historic development of events in order to react to trends.

3.3.1 Art as ambient displays

The use of art in information displays risks being seen as tampering with the aesthetic qualities of the artwork, especially if the integrity of the artwork is compromised. The subtle smile of the Mona Lisa could indeed become a broad grin in the event of a favorable increase at the stock market. But the artistic expression might suffer a downfall in the long run. The symbolic value of the components in a work of art will suffer from the problem of encoding even more than traditional icons do. The user would simply have to remember the reference before realizing the content. The bird could fly dangerously close to another element in the painting, as Stasko's team suggested (Stasko et al., 2004), but would that mean something?

The idea of creating collages out of artworks ignores the importance of the internal relations within a composition. With a digital painting, one can calculate the new positions of components accurately, but the resulting aesthetic composition of the components is unpredictable.

Holmquist and Skog recognized the growing ubiquity of displays and predicted the various shapes displays will take during this development: textile-displays, electronic inks (e-paper), and wallpaper (Holmquist and Skog, 2003). They suggested that this development would result in a change to the way we design interfaces, for example, from designing desktop interfaces to designing environments. This new role of designers taking the physical environment into consideration as architects or interior designers would brings with it new rules for how to deliver information.

Holmquist and Skog presented the concept of "Informative Art," artistically visualizing information in the environment. They developed prototypes from well-known artworks and styles that were especially adaptable for modularization. Modularized artworks lend themselves to a certain rule or method that easily may be described, transformed, and processed within a computational setting. The abstract compositions of Piet Mondrian consist of a rigid grid. Red, blue, and yellow rectangles of various sizes are contained within an irregular black grid. But instead of using an original work by Mondrian Holmquist and Skog respectfully copied his style for a visualization of global weather 3.14. The countries sourcing the temperatures are shown on the display as colored squares placed in approximate geographical alignment. The color of the squares follows the weather conditions at the various locations: blue signifies rain or snow, yellow clear weather, and red cloudy conditions. Furthermore, the size of a square was determined by the temperature at the relating country - the hotter the bigger.

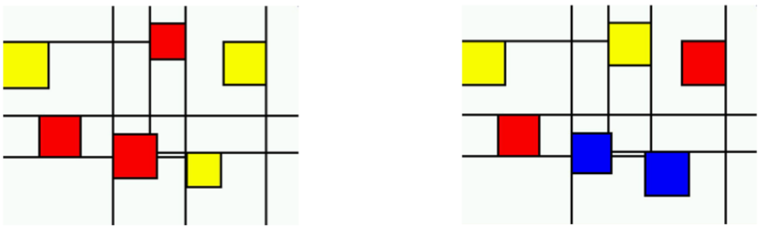


Figure 3.14: A Mondrian-inspired, dynamic "Weather Composition." The position of squares is determined by the relating geographical location of the referring city. Size is determined by temperature, and color by type of weather. Image (Holmquist and Skog, 2003).

Holmquist and Skog defined the informative role of Informative Arts as a less than exact form of communication. The subtle changes within the display and the abstractness of the mappings are better suited to an overall aesthetic impression of the information than to precisely readable measuring. Holmquist and Skog viewed the lack of visible change combined with the fuzziness of information as important means to maintaining a functional and aesthetic artwork. They insisted on a distinction between Informative Art and computer graphics:

> These are conscious design decisions that have been made to make the applications function more like a visually pleasing artwork, and less like animated computer graphics (Holmquist and Skog, 2003).

One could argue that this critique of dynamic artworks suffers from a conservative exclusive limitation in only seeing static paintings as possible means for wall-mounted visual communication. Holmquist and Skog claimed that changes to the display must remain hidden to the user in order to maintain calmness.

Presumably, Holmquist and Skog viewed Informative Arts primarily as a visualization method for nonessential information, and as such neither art nor information. The work of Holmquist and Skog provides interesting examples of the use of aesthetic displays to communicate complex information in a nonintrusive manner. Their approach is based on the assumption of one-way communication in which displays basically supply information (provided that the users know how to read the mappings) or simply serve as decoration.

3.3.2 Ambient displays on everyday objects

Ambient displays can function through different modalities: using sound, tactile cues, light, color, or even changes to temperature.

Schmidt and colleagues argued strongly that location awareness is insufficient for achieving a ubiquitous display (Schmidt, 2004). They suggested using displays on-location to supplement the limited displays of mobile devices. The unobtrusive push of information may be achieved by distributing information to displays or devices contextual to content. Schmidt's group formed this notion around what they identified as "decision points" at which information would be displayed timely. Addressable displays will be omnipresent in the future, they predicted. The umbrella stand could flash on a day with a forecast of bad weather.

Schmidt's group defined three design criteria that must be met when embedding information into the environment. 1) Information should be embedded where and when it is useful, that is, at "decision points" or at locations most contextually relevant to the information. 2) The design should take care to embed the information in the most unobtrusive way. It should ambiently provide the user with awareness of the information. No action should be required from the user. 3) The displays should be dedicated to the content of the information they deliver.

The idea proposed by Schmidt's team to avoid interaction with the user raised some concerns. If every display has to be dedicated to a single purpose, and if displays may be used at a level of detail at which even an umbrella stand has its own display, the result may easily become as noisy as Times Square. For a particular user, a dedicated display for an umbrella stand might have some value. It may coincide with his schedule for the day, and with his ownership of an umbrella in the stand. However, a multiuser-scale dedicated display would have to decide whom to address, or consequently deliver information out of context to some people. If persons A and B are at a decision point, but only person A is going out, should the umbrella stand light up? The resulting chatter of displays sometimes addressing people, sometimes not, could endanger the perceived value of the information and contradict the design ambition of calmness. If Schmidt and colleagues' intention was to go beyond location-as-context, the idea of assigning information displays to devices is still in line with location-as-context and suffers the same shortcomings.

Elliot and Greenberg presented various examples of displays resembling, or embedded in, everyday objects (Elliot and Greenberg, 2004). The "Glow Lamp" rotates a multicolored screen around a light bulb, changing the color of the light according to dynamic information 3.15. Elliot and Greenberg examined the possibility of combining displays with sensor technology to achieve a direct interaction. Their "Ambient Beads" consisted of colored beads on a string. The height of a bead signified an event or state, for example, "time 'til lunch" or a friend's availability. Each bead is equipped with a sensor, so touching a bead could connect to an action, for example, opening an instant-messenger window 3.16.

Figure 3.15: Glow Lamp by Elliot and Greenberg. The lamp can rotate its shade on its main axis to change the color of the light emitted. Image (Elliot and Greenberg, 2004).

A bead continually moving up or down may directly represent a linear time scale, but the information causing a lamp to change color isn't comprehensible unless the user remembers what a particular color represents. The approach to embed sensor technology into displays explored by Elliot and Greenberg may

Figure 3.16: Ambient Beads. Each bead is suspended by a thin wire from the top of the monitor, and a small motor controls the height of the bead by reeling the line in or out. A continuous motion, or motion at regular-intervals, may be mapped to an approaching event in time. The changed position of a bead could be mapped to a friend's online status: "up" or "down." Image (Elliot and Greenberg, 2004).

solve some of the issues concerning what events mean or represent, in that the user may get additional information from the display.

Recently France Telecom developed a low-resolution display called "Create Wear," which is sufficiently robust and flexible to embed into clothing. People use this "communicative clothing" to display information to the environment rather than receive it. The display is accessible via a mobile phone or a wireless network 3.17.

The design group at France Telecom has come up with a variety of uses for this technology, both practical and social. A back-mounted display functions as a blinking turn signal for bicyclists. Other displays express feelings like "I love you" 3.18.

The target customer group is attention-demanding young people. France Telecom suggested that this technology could be used to express statements, moods, and so on. The display could even be connected to sensor equipment that would trigger predefined animations and statements. Research and Development Project Manager Emeric Mourot said in a press release,

> Clothes are becoming a key interface for giving graphic expression

> and form to your moods. It's a very personal symbolism, an emotion or state of mind that you can now display publicly and very simply through eye-catching animated graphics and short texts (Telecom).

New, wearable, large-size displays like those from France Telecom could also be used to transmit messages to information systems equipped with optical sensors. A certain pattern could give the wearer privileges or access to certain areas.[1]

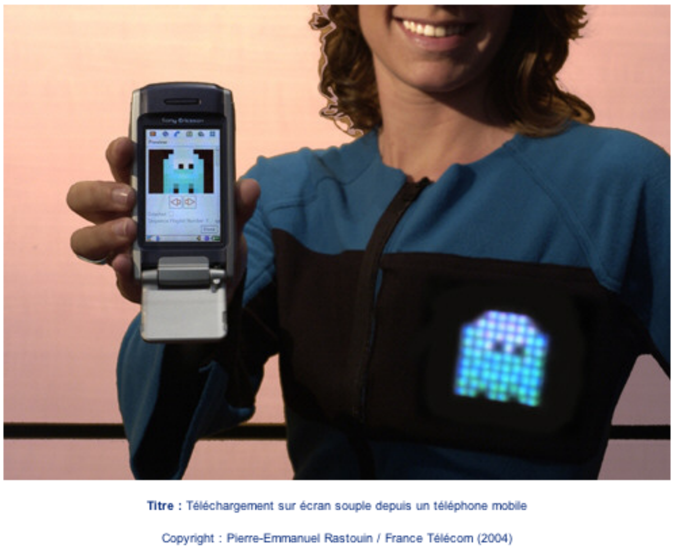


Figure 3.17: Wearable-display image by France Telecom being transferred from a mobile phone. Image Pierre-Emmanuel Rastoin, France Telecom.

3.4 Ambient projectors

Pingali and colleagues (Pingali et al., 2003) proposed the use of projectors as a means of delivering interfaces. The IBM Everywhere Displays Projector is an

[1]Patterns shown on the personal display could be like "Quick Response," developed by Sony. Sony has promoted its PlayStation Portable (PSP) platform using coarsely bit-mapped abstract patterns on posters. The patterns contained a visual code called Quick Response, which could be interpreted by mobile phone cameras that could translate the message into a URL. Quick Response patterns are becoming ubiquitous in Japan. The patterns need not be printed in large format to be readable by mobile phones but may reside simply in the corner of a poster.

Figure 3.18: Wearable displays by France Telecom. Image Pierre-Emmanuel Rastoin, France Telecom.

attempt to design a system for intelligent pervasive computing spaces for which the display is allowed to follow the user around in an environment (Pingali et al., 2003) (Sukaviriya et al., 2004). The system is composed of a projector combined with a computer-controlled mirror and camera 3.19. The resulting "steerable interface" projects displays and interfaces onto surfaces in the environment.

Pingali's team proposed the steerable interface as a suitable solution for inducing the interface to appear when and where the user needs it. The combination of sensors (the camera) with the display technique forms the basis for automatically delivering displays in the vicinity of the user. The camera also adds an interactive dimension to the display, enabling it to become an interface. By automatically triggering the appearance of the displays on the basis of user location, the content of the displays can vary according to the proximity of the user. Pingali and colleagues proposed shopping as a scenario for this type of function. A customer may be drawn to the appearance of information concerning a nearby product. As the customer closes in, the information may become more detailed or even become an interface for ordering the product.

The distinction between a display and an interface is accentuated in Pingali and colleagues' six-point definition of what constitutes a steerable interface. They proposed that a steerable interface should be capable of performing both input and output functions, as these support contextual and casual interaction. The lack of need for devices in order to interact with the system is made possible through the appearance of virtual keyboards, push-buttons, or the like, on surfaces related to the task at hand.

In order to maintain a high visual resolution for the conventional pointer interface used, the image has to be corrected to counter distortions caused by the angle of projection.

The system relies on "Environmental Modeling" through the use of 3D models and object descriptions (e.g., the optical properties of surfaces) and how they are best put to use. This information undoubtedly helps produce a higher degree of acuity to the system, and work is being done to generate anti-distortion automatically (Pinhanez, 2001). Even though anti-distortion is a salient feature of the system, the cost of achieving this may be high. By introducing a model of the physical environment, the system can become dependent on information that is not contextual, as the model would need constant revision. Automation of the revision process could solve nonalignment between the model and the real world by simply constituting a prerendered model with an interactive one. This solution would require the automation of counter-distortion and brightness settings by trying out surfaces dynamically using visual calibration patterns. The ongoing scans for suitable surfaces could be undertaken at a pace that is faster than human visual detection.

Pinhanez argued that the use of projecting techniques could help in guiding people around a physical space by indicating directions or constraints on the ground (Pinhanez, 2001). The Infineon and Vorwerk companies have literally made use of the ground as a display. They produced carpets with built-in sensors and LEDs that may be used to guide people around (Infineon). Many other uses have been proposed by the companies, including climate control and fire and burglar alarms, as the carpet sensors can measure temperature, pressure, and vibration 3.20.

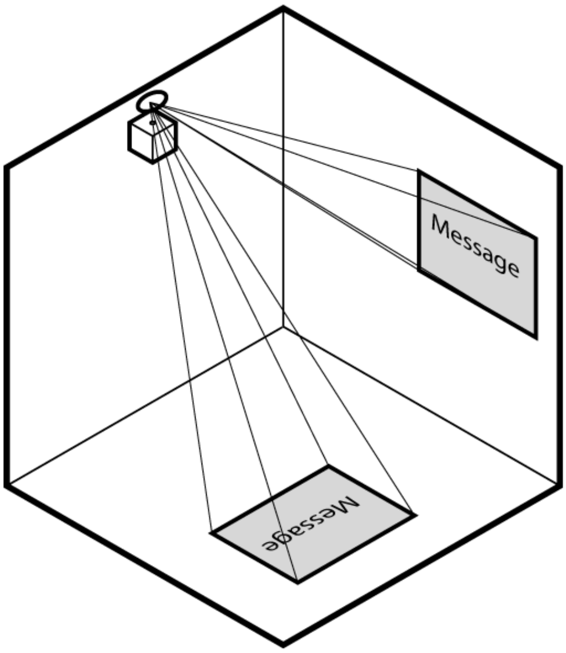


Figure 3.19: The Everywhere Displays projection system works by directing a projection with a mirror.

Figure 3.20: A carpet display can be used to guide users around. Image (Infineon).

The task of guiding people using projections could be achieved with lower resolution LCD-projectors. For instance, a Moving Head stage-light 3.21 could follow a user around, displaying specialized information at decision points.

Figure 3.21: A commercially available lamp like the MAC550 remotely controlled fixture from Martin Light could be used to project information. Image www.martin.com.

An inexpensive but powerful alternative to commercial projectors could be created from a cluster of laser pointers 3.22.[1] The assembly needed for an 8x8 pixel resolution would consist of 64 laser pointers, a control board, and some sort of connection to a computer, like a serial or ethernet.

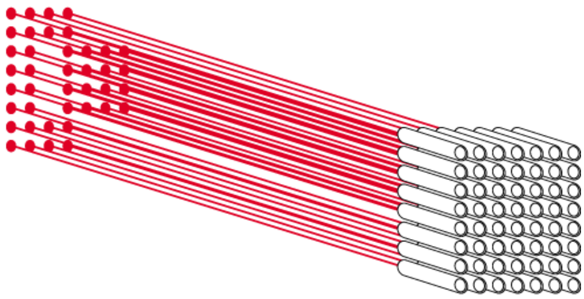


Figure 3.22: A battery of laser pointers could project low-resolution graphics across long distances.

[1]This solution would require taking caution to project only onto surfaces from positions not crossing the path of people, as shining lasers into people's eyes is not considered ambient.

3.5 The push of information

Protecting personal information from the prying eyes and ears of outsiders is a major design task when working with ubiquitous displays. Securing personal data has been less an interface problem than a security-through-encryption challenge for mobile devices due to the limited ways to communicate information. The limited screen size of a mobile phone is enough to prevent "eyes-dropping." The possibilities created from communicating by means of public displays present challenges in terms of privacy. Various ways to protect personal information in public communication has been proposed for instant messages that present information solely to the user by means of coding (Tarasewich and Campbell, 2004).

Using coding does partially solve the protection of the information content, but at the cost of information not being readily perceivable by the user. Codings may disguise the content displayed, but the context at which information appears could in itself be revealing. The reason for the appearance of messages could be interpreted or guessed at if the messages appear in response to environmental events. If a public announcement coincides with a personal display's coded message, users may notice the connection. In this way, a coded display could be as revealing as a dog's wagging tail.

To preserve ambience, people should not be required to use code books or translations to perceive information. Information should be readily available to the user. Instating codes or ambiguous signals prevents ambience, as people must remain attentive to the manner in which information is encapsulated rather than to the communicative content of that display.

Protecting personal communication is primarily a question of knowing what is the state of the information system. The information provider may contain various states of profanity, and the manner in which information is displayed should reveal that openness in a public announcement system. This could be achieved by tone of voice. A voice addressing you with an echo could mean that the service is talking about public matters, whereas a voice addressing you in a whisper could signify that personal matters are being conveyed. While this solution doesn't prevent a display from broadcasting matters, the user readily recognizes that the system could convey delicate information. If he felt uncomfortable with the situation, he could choose to silence the system.

How do you send a private message to a person via a public display? Tarasewich and Campbell did research into pixel-based displays that used colored lights instead of text or images to display information (Tarasewich and Campbell, 2004).

The advantages to using minimal display techniques (i.e., three-LED or three-pixel displays) lie in the portability of the display. The very small size ensures that existing artifacts can adopt the display and augment its use. A ring could blink at an increasing interval to communicate an upcoming meeting appointment. Furthermore, the minimal power consumption of such displays makes them well fitted for ubiquitous use.

Tarasewich and Campbell argued that the high degree of decoding needed to read their three-pixel display acted as encryption to bystanders. An LED on a finger-ring that switches betweed purble, yellow, and red would present very little information to anyone but the owner of the ring. Even though the signal is difficult to read, it is not protected by encryption. Tarasewich and Campbell acknowledged the shortcomings of limited visual displays to convey information. They estimated that, theoretically, a high number of messages can be created using three-pixel displays, but they expected 8-11 different pixels would be more practical.

Messages could be sent from your car, alerting you that the parking meter is about to expire or that someone is trying to break into it

The problem with three-pixel displays seems to be the lack of variation in the messages. Even though each LED supports two brightness levels, the system isn't suited for dynamic or multiple messages. Your spouse may have tried to call you, indicated perhaps by a single red light, but the content of the call is unknown. The system quickly runs out of easily perceivable combinations of visual patterns. Complex messages could indeed be sent to a minimal system as Morse code, but this requires focused attention from the user in order to decode. Decoding Morse is indeed a complex skill that needs specialized training to master.

An attempt to solve the problems of intrusiveness and privacy in relation to the use of mobile phones was undertaken by Nelson and colleagues (Nelson et al., 2001). The primary focus of their work was to be able "to respond to telephone conversations without talking aloud." The approach for their work to mix the modes of communication by combining voicemail with conventional telephone communication. The system implements prerecorded or predefined voice messages, which can be manipulated by the user through a three-button PDA or phone interface. In this way, the caller can interact with the user without changing modality while the user manipulates a graphical user interface. The graphical user interface for the system was created with low-level attention in mind. Instead of a complex menu structure, the system interface consists of only the three buttons.

Nonintrusiveness is important to ambient use of technology. In a bus or any other public place, most people would probably appreciate ring tones to be replaced by visual cues sent only to the owner of the telephone. However, if you are engaged in a face-to-face discussion with your colleague, who then suddenly excuses himself and starts talking to the air, you would realize that he had not been paying attention to you for the last 2-3 seconds. Yet, if his telephone rang, you would stop talking immediately, realizing that he was unavailable for the time it takes to either turn off the phone or answer the call.

Cultural habits are subject to convention, and behaviors may change. The private nature of head-worn displays hold social caveats, in that the interpretation of attention level and actions between people may become difficult. Billinghurst and Starner acknowledged the risk of social isolation in relation to wearable computers (Billinghurst and Starner, 1999). Normal rules of conversation become hard to follow if people cannot estimate each other's engagements. It is a common mistake to respond to the vocal outburst from people wearing telephone headsets. The "nonobvious" activities associated with head-worn displays carry the danger of making them socially unacceptable because they create uncertainty about the user's attention level.

Starner related an example of an interview in which the reporter maintained eye contact with a subject using a head-worn display for note-taking. In this context, use of the head-worn display may be considered acceptable, due to the special social roles associated with an interview (Starner, 2003). Other situations may be influenced negatively by the use of nonobvious interaction technologies.

Similarly, the experience of talking on the phone with someone who is engaged in other matters, like checking email, can be very annoying.

If the use of head-up displays cannot incorporate clear indications of the user's intentions and actions visible to other users, this takes its toll on the human-to-human interface, resulting in invalid or unclear facial expressions. The user of the head-worn display may run the risk of being misunderstood due to his distributed attention and nonobvious facial expression. Emotions play an essential role in human intelligence, cognition, and perception (Picard, 1995). The possible effects of technology obscuring human-to-human interaction could prove disastrous to the use of mobile devices in situations of social interaction. The use of mobile devices expands the interaction and interface concerns to the physical environment and other people. The introduction of mobile devices should not inhibit the qualities of natural situations, such as by removing or diminishing a user's facial expressions.

Experience from the desktop computer has taught us that attention-grabbing means of communication (flashing colors, animation, pop-up windows, etc.) should be used with care so as not to confuse or distract the user.

The multitude of ubiquitous information displays revived problems with distraction associated with commercial web sites. Banner ads often compete strongly for our attention, which we then may experience as annoying or distracting. This is not what we want to experience while looking at the navigation system for our car, or when using our PDA to check for new messages at the local coffee shop.

However, ignoring the digital world in the same way we ignore banner ads would be a great loss also. I may want my navigation system to send me a message that I am approaching a place to eat lunch that suits my preferences. I may want to know that Starbucks has a network printer so I can pick up that document my colleague just mailed me. The question is, how to allow for changing services to address me while I continue working on my primary task of interest. How can one be driving to work and yet notice that a parking space is available out of the corner of your eye? Should we allow the parking meter to indicate the spot's vacancy, and if so, how do we avoid the intrusive "paper clip" effect?

The incoming messages should be balanced perceptually, so as to be noticeable yet not take focus away from the primary task. On the formal level, we want to filter the incoming messages just as we filter spam email according to our preset preferences. The messages themselves could contain meta-information as to the level of privacy, indicated by the sender, just as we may indicate level of importance on email.

3.6 Summary of new interfaces

I have now presented a selection of research that uses post-WIMP interfaces to expand interaction into dynamic environments.

The major challenge of these new displays is related to the attention capabilities of the user. How can the displays be peripheral yet detectable, and as for the concern for privacy, how can they be delivered in public, yet be private in an uncoded way?

Sound seems to be a particularly strong modality to meet these requirements, as sound by nature has ambient, peripheral qualities. But visual displays may also work well for ambience, provided that they are discreet yet detectable. Detectability in terms of both their visibility and also the user's being able to recognize that a message is meant for him - not for some of the other people present in the public space.

The last part of this thesis takes these points to the design of a specific interface for ambient intelligent services, the DELCA project.

A major source for inspiration stems form the basic visual element that digital media provide - the pixel. I therefore first examine the matter of pixels.

Chapter 4

Low resolution

In the previous chapter, I argued that design formats for ambient displays should be robust against changes in resolution and perspective and against occlusion. In this chapter, I argue that this robustness may be achieved with low-resolution format. I present some outstanding works of art and design that demonstrate how low resolution can meet the challenges without sacrificing aesthetics. Therefore, I end this chapter by suggesting that low resolution may become the common style for the first generation of ambient visualization - just as it has now become a signifier of first-generation computer games.[1]

In modern days, we see consumer department stores using amazing plasma-screen displays with incredibly high resolution. The quest for photo realism as the solution to technical and visual needs has been underway for some time. However, with respect to the design of ubiquitous displays, the craving for high definition to compose reality may drive us into a dead end. In this chapter, I hope to draw attention to the historic development that led to almost photorealistic high definition on the PC screen. I point out important landmarks in this development as a source of inspiration to the possible design of ubiquitous displays, and argue

[1]My personal experience with pixels started in 1983 with computer games on my trusty BBC Model B. In Basic "graphics mode," resolution was easily shifted from fine lettering to coarse blocks of light. The blocks were conveniently used in games, serving as monsters, spaceships, and laser beams. Never mind that a laser beam was ridiculously large compared with a spaceship, pixels were adaptable. Later in the 80s, I had the opportunity to work with the Quantel Paintbox and its revolutionary possibilities; an attached video cam could capture the analog world and let you work on the resulting bitmap image. Despite the fact that resolution was extremely high at the time, the zoom function could still fill the screen with big square pixels in wonderful harmonic variations of color.

that the obsession with accuracy of detail may lead to the overall picture getting blurred - so to speak. I explore and draw inspiration from some of the new developments that do not base themselves solely on high-resolution displays. The main purpose of this chapter is to inspire the reader or designer to rethink how visual displays might look, and especially to recognize the potentials for motion graphics to present causal relations.

4.1 Grid formats

From prehistoric hand painting in caves to polygonal 3D constructions, the tools for expression have defined the messages. The application of a system to expression has preset limits for the expressional possibilities. For example, in carpet weaving the number of strings and size of the weaver's hands set limits to the resolution.

In mosaics the regularity of the grid is adjusted for the individual pebbles to form precise shapes (4.1).

Figure 4.1: Detail of mosaic from Piazza Amerina, Italy.

The dominant technique for digital visualization is by the use of pixels. The number of pixels is defined by the monitor's resolution. The resolution of a screen forms a Cartesian grid that defines the pixels.[1]. The grid is what defines

[1]Pixel n : (computer science) the smallest discrete component of an image or picture on a CRT screen (usually a colored dot); "the greater the number of pixels per inch the greater the resolution" [syn: pel, picture element] (WORDNET, 1997)

our screen space.

4.1.1 Technique

Various techniques have been developed to counter the limitations of color range and resolution in static raster graphics. This need for flexibility within the rigid system of color and resolution has spawned aesthetic solutions now considered classics (4.4).

Raster graphics are defined by a rigid grid, which in the static mode has to turn to "anti-aliasing" to form shapes that are not readily shown by the grid. Without this loss in visual definition "jaggies" would appear. Anti-aliasing solved the basic problems of displaying detail in raster graphics experienced when rendering (Crow, 1977):

- Edges of an object
- Very small objects
- Complicated detail

When displaying small text, the blurring caused by anti-aliasing is particularly noticeable 4.2.[1] A reaction to counter this phenomenon could be seen in the late 90s with the surge in popularity of bitmap fonts and bitmap art in general among web designers. The minimal bitmap fonts are perfectly suited to small sizes due to their aliased nature that supports visual clarity. One of the most popular is Jason Kottke's Silkscreen font 4.3.

Crow argued that small objects pose a problem when displayed aliased because they may disappear when scaled or displayed from different angles (Crow, 1977). A thin line could be transformed at angles into a "string of beads" because the angle of the line comes into conflict with the resolution of the grid 4.5. Despite their unnatural appearance, these jagged and isolated dots and lines have been generally accepted in the everyday use of technology, as with emoticons like "smileys." These small articulated faces hold a great deal of abstraction, and still they provide meaningful information to users.

[1]Please note that graphics programs have become much better at solving this problem nowadays. Adobe Photoshop offers many different anti-aliasing text styles, among them sharp, crisp, smooth, and strong.

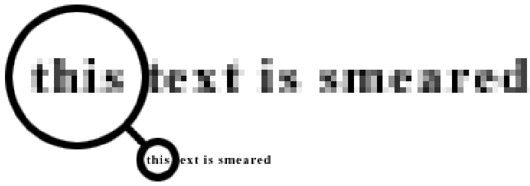


Figure 4.2: Text smeared by anti-aliasing.

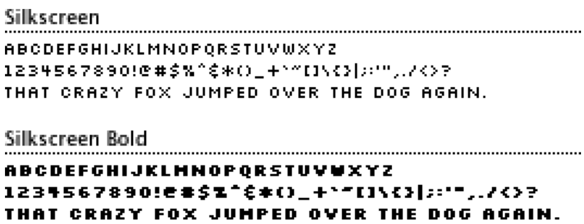


Figure 4.3: Kottke's Silkscreen bitmap font.

Figure 4.4: Screenshot of Susan Kare's "Woodcut" in MacPaint 1983. Image http://kare.com.

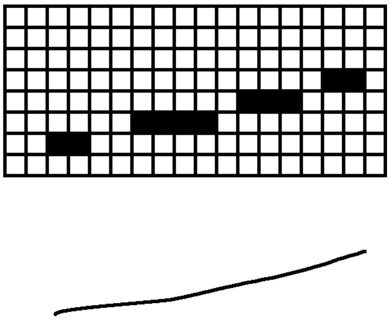


Figure 4.5: Line-to-pixel conversion. Without anti-aliasing, lines at angles close to the grid lines may result in the "beads" effect.

Computer graphics designers realized that anti-aliasing only works well with images of high resolution. If the image area becomes too small, the anti-aliasing blurs the motif and destroys visual acuity 4.6.

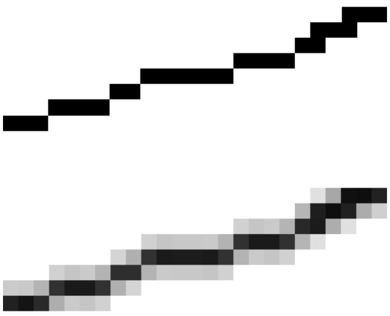


Figure 4.6: Closeup of aliasing versus anti-aliasing effect. Top: aliased line. Bottom: anti-aliased line.

Raster graphics are commonly indexed. This means that the color range is limited. In early web development, many monitors could only show 256 colors. That meant that once the system had reserved colors for system use, a browser was left with the remaining palette. Using colors outside the 256 colors already in use by the system would result in a replacement by the system to a color found in the palette. If the palette consisted of 256 shades of green introducing an indexed image of a sunset, the system would display the image using green colors.[1]

The development of indexed graphics span new methods for controlling transformations of color outside the reach of the limited palette. Producing gradients used to require optical blending of color, or displaying a figure with seamless transitions between multiple colors or light values, as in the example of adding green color to turn red or black to white. Such blending techniques utilized "dithering" to even out the distribution of pixels by swapping them around 4.7. Indexing colors is yet another way of imposing a rigid structure on visual media, which in turn has shaped the aesthetic of digital media. I focus on monochrome raster graphics in the design for ubiquitous displays, because the lowest common denominator is black and white, as seen on low-cost small-size watch displays 1.1.

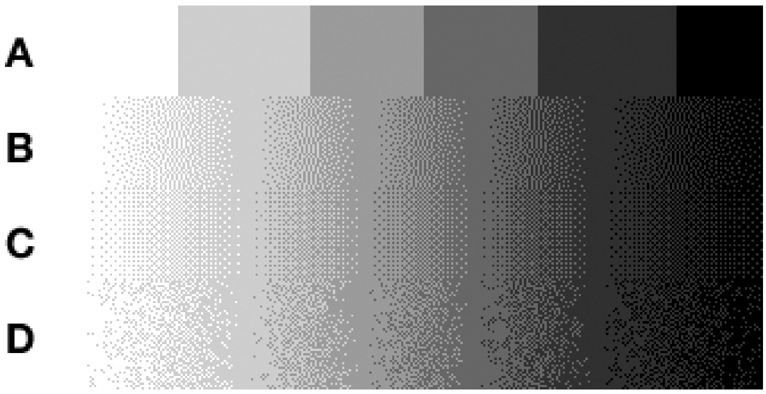


Figure 4.7: Example of dithering techniques. A: no dither, B: diffusion, C: pattern, and D: noise.

4.2 Nonhomogenous grid displays

Our notion of the pixel is built around a regular grid. However, the irregular grid is a particularly interesting alternative for displays in the "wild" physical environment. When we move outside the framing of the conventional computer

[1]The web 216 is an industry standard defined by Netscape's 216 "websafe" color palette, meaning colors that would be indexed to their true color value. If we judge from ubiquity, then the web 216 palette of the 20th century has had a greater cultural impact than the 19th-century discovery of artificial dye techniques. Even now, long after monitors ceased being limited to 256 colors, designers are still rooted in the web 216 palette.

monitor, a display can occur in a multitude of ways. We will encounter displays varying from small monochrome ticker displays to gigantic color screens. At times, we may stand directly in front of the displays, and at other times, parts of the displays will be occluded. An image projected onto a surface may be compromised by the topology of that surface. We may even foresee new types of displays woven into clothes that will be affected by the folding and curving resulting from the wearer's movements.

One common irregular-grid display is "text television." In this system, shapes are displayed using nonuniform pixels in a 2x3 rectangular pattern rather than a 1x1 square.

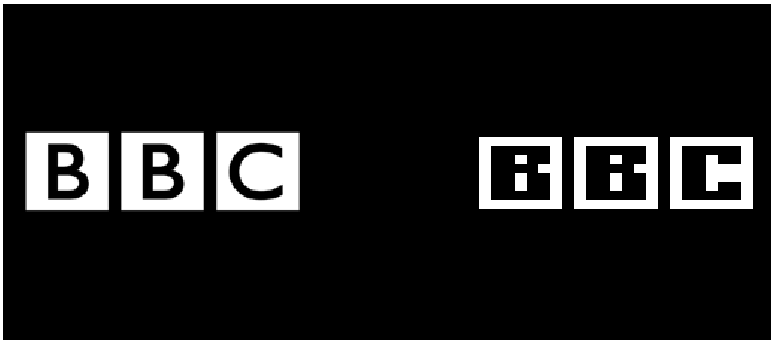


Figure 4.8: Left: BBC logo. Right: The BBC logo converted to the the BBC CEEFAX Teletext system grid.

Heaton explored irregular grids with nonuniform distribution of pixels. In the Nami project she created a novel display system that fundamentally challenges our notion of raster graphics. She developed a series of autonomous 1x1 pixel displays with wireless communication capabilities using bi-directional infrared (IR) sensors. Nami is constructed around decentralized networks, and each display or physical node[1] is not aware of the spatial location of other displays in the network. However, using a related interface called "Digital Palette," Heaton was able to share images across the network and have each display node portray a portion of the image while still keeping the structure of the image perceivable 4.9. The resolution of Nami is extremely low, but nevertheless the principle of

[1] In the actual prototype, the communication is done by IR links.

ad-hoc composite displays seems to work, and the system offers an interesting insight into nonuniform or amorphous display systems (Heaton, 2000).

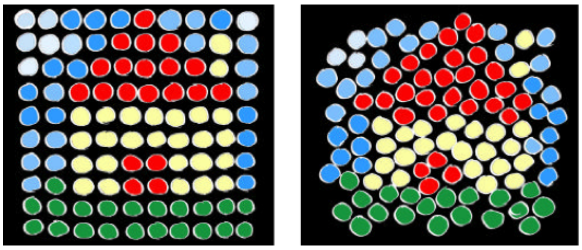


Figure 4.9: The two grid systems. Left: conventional. Right: the Nami amorphous grid. Image Heaton.

If we have to leave the "safe" regularity and frontal perspective of the Cartesian grid, we challenge fundamental communicative aspects like the recognition of shape. When the grid of the ad-hoc network changes, so too does the image portrayed. The unstable resolution of the display requires that the characteristic properties that define an image (i.e., "invariants" according to Gibsonian terminology (Gibson, 1979)) be robust enough to survive unpredictable visual transformations.

Ensuring robustness in an image requires attention to detail. But images dependent on detail run the risk of becoming ambiguous when shown at various resolutions. At low resolution, a detailed image will lose some information, and at changing resolutions, the same image could lose visual significance 4.10.

4.3 Motion at low resolution

Surprisingly, motion has not been widely employed in interface design. A few years back, restrictive bandwidth and processing power could be blamed for preventing the use of motion graphics, but these factors have now become less of a problem.[1] In cellular automata like John Conway's "Game of Life," simple

[1]In *Being Digital,* Nicholas Negroponte argued: "One of the reasons you have seen so few dynamic graphics or you have found video displayed in a small window is that it is hard to get sufficient pixels fast enough (to produce 60 to 90 frames a second needed for flicker-free smooth motion). As each day passes, somebody offers a new product or technique to do this faster" Negroponte (1995) (p. 106).

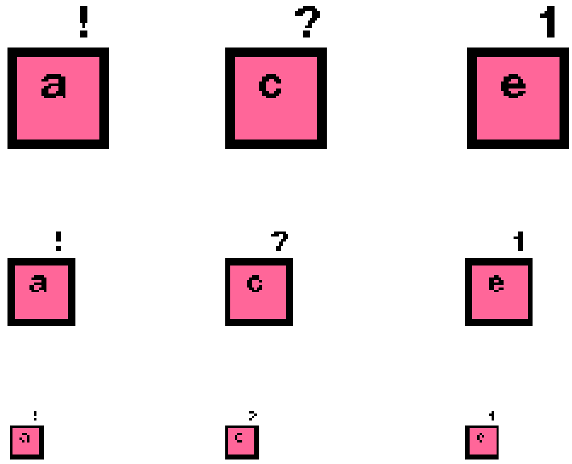


Figure 4.10: Three poorly designed symbols shown at different sizes. In the smallest version, the symbols become difficult to tell apart because of the loss of detail.

rule-based functions that resemble cellular growth form coherent visual structures (Gardner, 1970). The visual structures not only exhibit self-maintenance within the system, but do so in highly recognizable ways. The coarse automata patterns that emerge even have names referring to their animated appearance: "Traffic lights," "Unix", "Two eaters," "Small fish," and so on 4.11.

Figure 4.11: Game of Life figures, from left: "Traffic lights," "Unix," "Two eaters," and "Small fish."

Motion is a promising method for handling the challenges presented by low-resolution displays. Gerda Smets argued that moving patterns of extremely low resolution (e.g., 21x12 pixels) may be recognized as easily as images of high resolution (Smets, 1995). Smets maintained that the perception of flow in moving patterns forms a precise impression of the information 4.12.

Figure 4.12: A single frame from Gerda Smets' "flying bird" at low resolution (Smets, 1995). When all the frames are played as a film, an elegant eagle flies.

The use of motion in everyday life is essential not only to our physical displacement but also to our perception and reasoning. Motion is a promising candidate for representing information that doesn't need decoding. We perceive the world by movement (Gibson, 1979):

> Observation implies movement, that is, locomotion with reference to the rigid environment, because all observers are animals and all animals are mobile.

Motion is related to living things.[1] Therefore, movement is a good way to display events and states, as these are dynamically instated. Autonomous movement is generally conceived of as relating to living things, whereas blinking is often related to mechanical or digital apparatus.

Motion is a potent way to attract users' attention. Motion can be mentally economical, due to its preattentive and interpretational properties (Bartram, 1998). Bartram investigated the use of animation in the periphery, and found that the distraction is dependent on the motion pattern applied.

Motion also presents a preattentive level of grouping, which adds to the perception of complex structures (Bartram and Ware, 2002), and, in the case of our

[1]In *The Child's Perception of the World,* chapter 6, Piaget describes interviews with children who define life as being associated with movement. Life is a "continuum of free forces endowed with activity and purpose..." (p. 206) (Piaget, 1929).

figure design, adds coherency to fragmented figures such as composite body mass (see p. 92).

Using motion correctly by avoiding unnecessary distraction is, of course, essential for a pleasant result. Studies have shown that an optimal balance between a primary task and a secondary source of information can be achieved by careful design. D. Scott McCrickard explored scroll and fade effects in dynamic displays, with promising results (McCrickard, 2000).

Robertson and colleagues emphasized the smoothness of animation combined with interaction as a good way to lower the cognitive task requirements:

> Smooth animation is particularly important because it can shift a user's task from cognitive to perceptual activity, freeing cognitive processing capacity for application tasks. For example, interactive animation supports object constancy (Robertson et al., 1993).

In general, the use of color, shape, or symbols in the periphery of the visual field is hard to notice because of low acuity. Low-resolution displays set limits to the use of text or symbols. Motion allows for the display of additional information in a limited area thanks to the temporal dimension.

Humans can perceive relational properties, like causality, even with simple, low-resolution graphics. Causality contains latent perceptual levels we have come to take for granted, without realizing the potential of this phenomenon. In 1946 psychologist Albert Michotte first reported a series of experiments on the perception of causality. He devised an elaborate mechanical apparatus that allowed him to manipulate the animation of two objects on the projection screen 4.13.

These animations could be as simple as moving a dot at various speeds and with various delays. However, small movement variations in the display would result in significant variations in the description of what his subjects reported seeing. Their reports might vary from the factual descriptions,"A launches B," to descriptions that contained words attributing motivation, emotion, gender, age, and relationships to the two objects:

> It is as if A's approach frightened B, and B ran away (Michotte, 1963).

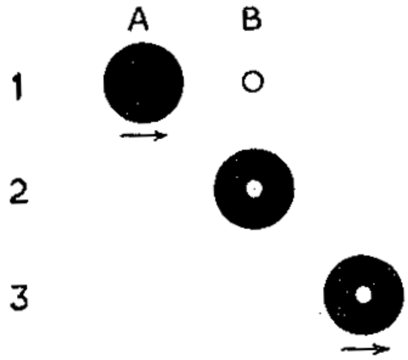


Figure 4.13: One of Michotte's schematics of a visual event. The circle B starts to move with A when they are concentric. Image from (Michotte, 1963).

The invariant properties capable of unleashing a sense of causality in us was found to be temporal:

> The impression of causality is dependent on specific and narrowly limited spatial-temporal features of the event observed. The spatio-temporal organization is such that it directly unleashes this impression in us. Alter the relevant variables by a small but measurable amount and the impression disappears (Oldfield, 1963) (In (Michotte, 1963)).

Movement may have the potential for escaping encoding or even objectification. In Michotte's discoveries concerning object-removal, he wrote:

> It follows that the movement of an object is liable in some circumstances to survive phenomenally the removal of the object, and there can be an apparent continuation of the movement of the object which has ceased to exist and is consequently not itself being displaced (Michotte, 1963).

In some cases, subjects would perceive causal relations between dots - even in cases where only one dot was visible at the time of the impression. This phenomenon could prove very valuable in bridging the physical gaps between displays in an environment. If a pattern is shown on display A, but the movement of the user makes display B more attractive, then the timing at which the image leaves screen A and appears in screen B is essential for the user to maintain the experience of object persistence[1].

The use of animation has also been examined in ecological psychology, perhaps most notably in Gunnar Johansson's point-light experiments. Actors wearing small lights all over their body performed actions in dark rooms (Johansson et al., 1994). The actions were unrecognizable when the actors were standing still. However, when the actors were in motion, the groupings and relations among the lights made it possible for the subjects to perceive gender, weight, type of action performed, and even the moods of the actors. These ecological experiments point to interesting relations between our perception of movement and personal identifying traits such as gender. In our design, we have not yet been able to draw upon these relations, as further studies are necessary to identify these invariants at low resolution.

The use of animation in communicating abstract relations was accentuated by R. Baecker and D. Shermann, who visualized sorting algorithms. Their claim was that a student who had watched the 30-minute animation carefully would be capable of programming some of the sorting algorithms. Although this was never documented or tested, interviews with the students supported the use of animation to communicate concepts. Baecker and Shermann explained why:

> The film also goes beyond a step-by-step presentation of the algorithms, communicating an understanding of them as dynamic processes. We can see the programs in process, running, and we therefore see the algorithms in new and unexpected ways. We see sorting waves ripple through the data. We see data reorganize itself as if it had a life of its own. These views produce new understandings which are difficult to express in words.

The possibility of creating taxonomy of movement could present us with a preattentive nonencoding approach of visualizing information. Paradoxical cases such as Michotte's outlined situations in which experimental events would cause an impression but in which real-life events would not represent a field laid out to us to be explored in the context of communicational structures.

[1] The CD-ROM contains a small artistic experiment with a causality effect: "Michottian".

4.4 Mechanics and aesthetics of low resolution

The restrictions found in the visual structure of computational systems have inspired artistic explorations that have formed a whole new style of digital aesthetics. One of the first artists to work in this field was the Korean artist Nam June Paik. In his "Magnet TV" piece from 1965, a television image was transformed by a powerful magnet positioned on top of the monitor 4.14. The magnetic power influenced the directional aim of the cathode ray tube (CRT) yolk, which resulted in the electron beam drawing a warped image. Paik managed to interact with the very process of building the electronic image in a very simple but elegant way. By manipulating the very path of the electrons composing the phosphorized image, the physical bonding of electronic media becomes evident. Paik played around with the rigid grid structure of the CRT screen and used physical circumstances in the environment to create an aesthetic transformation of the visual outcome.

Figure 4.14: Magnet TV, 1965. Television and magnet; black-and-white, silent; variable dimensions. Whitney Museum of American Art, New York, Purchase, with funds from Dieter Rosenkranz 86.60a-b. The Estate of Peter Moore/VAGA, NYC.

Daniel Rozin also does experiments with alternated pixel-based physical displays. His "Wooden Mirror" art installation mixes digital and analog forms in a display that uses 830 wood chips as pixels 4.15. The system operates by transforming color input from a camera into 35x29 pixel resolution in grayscale. Each wooden chip is then mechanically tilted by servos to reflect an overhead light

4.16. The degree of tilt applied to the wood chip is composed by a mixture of the individual chip's ability to reflect light and the digital gray value it portrays. The display is updated at 15 frames per second, resulting in both visual and auditory experience when the motors moving the chips spring into action.

Figure 4.15: Picture of the Wooden Mirror as it appeared in *Wired* magazine, Dec. 1999.

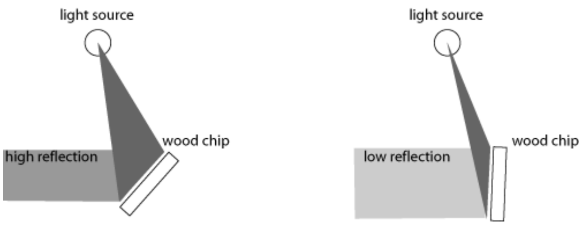


Figure 4.16: Principle of reflecting light from tilted wood chips.

On a macro scale, the architect Renzo Piano has worked with low-resolution graphics to add information layers to his buildings. The KPN building in Rotterdam sports a 3600-square-meter display system made out of 896 specially manufactured 24-volt light bulbs 4.17. Evenly distributed across the faade, the lamps act as a huge billboard that may be seen as far as two kilometers away. The minimal animations portray the nearby bridge or relay messages.

Figure 4.17: Left: KPN building by Renzo Piano. Right: closeup of display system, with additional standard ticker-style display.

Minimal displays using a limited number of LEDs has been extensively explored by John Maeda. His approach to design has inspired numerous researchers within the field of digital aesthetics. One of Maeda's latest works is the "ITU Interactive Light Walls" that consists of seven large LED displays in a 10x11-pixel resolution 4.18. The Interactive Light Walls are updated through a web interface. Four of them compose the "Two-chat" installation, which displays text messages, and the three red Interactive Light Walls display "Dial-an-art" abstractions on a number that can be sent to the display via the internet.

Maeda pioneered the rediscovery of low-resolution imagery through his Design By Numbers (DBN) programming language (Maeda, 1999). Maeda was motivated to create DBN because a number of his students were only familiar with commercial digital graphics tools.

In a 100x100-pixel format, the DBN creates an exploration and learning tool by directly coupling lines of code with the visual outcome of the produced code. This retro approach to computer graphics facilitates a logical understanding of how a computer processes information. DBN demonstrates that it is indeed

Figure 4.18: The ITU Interactive Light Walls, by John Maeda, in action.

possible to rethink visual communication in terms of resolution and methods of deployment. The close connection between executed code and visual result gives rise to unusual aesthetic solutions that follow generative art more than static graphics.

Maeda's students have followed in his footsteps with a number of elaborate projects that use low resolution to achieve new methods for communication. Among these are Nikita Pashenkov's "battery-sized lantern" [shown in (Maeda, 2004), p. 166]. As the title implies, it is a tiny portable display. The "pocketable low-resolution video display" by Albreight and colleagues [shown in (Maeda, 2004), p. 167) is capable of displaying short, low-resolution video clips.

4.5 Summary

We have seen how low resolution has pioneered visual communication and aesthetics on the digital platform. Low resolution may overcome limitations set up by conventional tools. Minimal graphics play along well with communication in physical space because their technological requirements are modest. Artists and designers have demonstrated how visual information in low-resolution displays can be as vivid and engaging as high-resolution displays. Perceptual psychologists have provided experimental evidence that people may perceive highly complex phenomena like causality from a very few elements in motion.

Chapter 5

Designing ambient agents

The DELCA[1] environment offers digital services in the form of "Ghosts." A Ghost may reside in a physical area defined by virtual information. They may interact with the user in the physical setting by mobile-device speakers, displays, or networked interiors like doors or elevators. They are aware of user-positioning data, and may themselves be positioned by the main system (e.g., not to leave a certain area).

The DELCA Ghosts are basically thought of as low-capacity services within an ambient intelligent system. If a user is talking to a Ghost residing around a water post, it is likely that the user will presume that the post Ghost will have more knowledge about refreshing drinks than, say, the stock market. The assigning roles to services explains the low cognitive ability of the individual system components.

5.1 The concept of DELCA

A prototype of DELCA is currently being implemented at the IT University of Copenhagen (ITU), which makes use of WLAN tracking, PDAs and other mobile devices, AIML (Artificial Intelligence ML), and an ensemble of more than 30 synthetic voices, each representing a Ghost[2] with individual specialties and

[1] DisEmbodied Location-specific Conversational Agents

[2] ghost n 1: a mental representation of some haunting experience; "he looked like he had seen a ghost"; "it aroused specters from his past" [syn: shade, spook, wraith, specter, spectre] 2: a writer who gives the credit of authorship to someone else [syn: ghostwriter] 3: the visible disembodied soul of a dead person 4: a suggestion of some quality; "there was a ghost of sarcasm

character traits. The system consists of audio, mobile displays, and ambient displays (Sørensen, 2004).

> Mrs. Jones enters the ITU a bit late for a meeting with Mr. Hansen. In the reception area, she is greeted by a male voice: "Hello, I am the Butler. May I offer my assistance? Please turn on your PDA." Mrs. Jones accepts the DELCA Ghost client on her PDA; it immediately recognizes the invitation for the meeting Mr. Hansen sent her the day before. The Butler continues: "Allow me to guide you to room 2.31, where Mr. Hansen will be joining us. Let us take the stairs to the left".

5.2 Ghosts

The social ladder of Ghosts ensures they can be ranked to a certain reasonable degree based on context. In this way, a user interaction with a formally inferior-ranked Ghost can be reinforced by the property of location to protect the Ghost from being chased away or interrupted by higher ranking Ghosts. This method of using location in combination with interactions to filter out unwanted information may be of great value to the usability of context-aware computing, in that the user interactions, to some extent, direct the flow of information.

A Ghost whose activities result in the achievement of user interaction is briefly ranked higher, as a product of its own formal ranking in the Ghost hierarchy and the added value of user contact. This results in a primitive brand of social intelligence being added to the Ghosts in that they cannot interrupt ongoing activities of combined higher rankings, but must wait for a more convenient time or choose a more discrete manifestation. The idea of dynamically tuning the manifestation level of Ghosts to that of their environment may help to develop communication techniques, with which information technology may handle larger amounts of data than currently possible.

in his tone"; "he detected a ghost of a smile on her face" [syn: touch, trace] v 1: move like a ghost; "The masked men ghosted across the moonlit yard" 2: haunt like a ghost; pursue; "She is haunted by her fear of illness" [syn: haunt, obsess] 3: write for someone else; "How many books have you ghostwritten so far?" [syn: ghostwrite] WORDNET (1997).

The location-awareness ability of the DELCA system enables interactions mediated by the environment. The Semaphore system design of distributing the user hierarchy according to location (Burrell et al., 2000) is similar to the DELCA Ghost design for developing location Ghosts. Location Ghosts reside within a certain area and may have special powers connected to that location. The Ghost "Printer Jan" can print out documents at locations and times convenient to the user. This feature is a functional aspect that shares the print example by J. Pascoe in the Stick-e note project (Pascoe, 1997), but it goes beyond the passive display of signs. Printer Jan is more than a preset sign that triggers an event; interacting with Jan causes position to becomes a function of the interface that may be actively engaged; interacting with the service fine-tunes the outcome.

The DELCA method of devising a variety of Ghost characters makes it easy to conceive entities for allocated functions in designated areas. Instead of accessing the DELCA system as a massive selection of multiple features, the system is divided into recognizable services performed by Ghost characters, which each exist at locations or situations contextual to the contained service. In this way the navigation within the system hierarchy is kept at a minimum, as the available programs or services change with the environment and user actions.

Ghosts act as mediators between the ubiquitous, omnipresent, transparent interface that is the user's environment and the specific situated-information spaces. Ghost characters may act as personal information lenses yet still be present throughout the environment; although the interaction principle of Ghosts is a shared one, the capacities vary.

I wish to propose a visual design that functions within this hybrid space of experience. One of the challenges related to the hybrid space is the unforeseeable number of events taking place within that dynamic environment. How to distinguish the service currently active from several others that wish to become active? A person may be interacting with one service while other services must wait with urgent messages they need to deliver to that person. How do we know which service we are interacting with, and how do we know that other services don't have a message for us?

The intention behind the conceptual foundation of agents as Ghosts is to transgress the boundaries between virtual and physical and play around with the narrative properties of an invisible disembodied being. This approach is radically different from the majority of agent visualizations in that the visual manifestation adopted by the Ghost is not to imply physical form but rather the activities they are involved in. Instead of replicating bodily features, the Ghosts' visual design should support the attachment of specific (synthetic) voices to specific invisible characters. The aim of my design presented in this chapter is to let low resolution signify the primitive state of the agents. In this way, the design doesn't try to emulate the appearance or capability of a human being, but constantly reminds the user that he is dealing with computer intelligence, that is, "a very fast idiot" (Tognazzini, 1993).

The virtues of low resolution were presented in the previous chapter. My design solution additionally favors low-resolution displays not because they may contain a representation of the world, but because they may relay the effects of hybrid space interactions and do so by virtue of low-cost distribution.

5.3 Visual format

The basic visualization of an available or detected Ghost consists of an 8x8-pixel animated pattern. On the PDA (IPAQ) the animation can vary in intensity as well as placement within a 10-pixel margin of the 320x240-pixel screen, thus leaving the majority of the screen exclusively for the primary task.

The unique pattern of a Ghost signifies which type of Ghost is present. The animation cycles of the pattern may be linked to, among other things, the activity state of the Ghost or the current strength of wireless signal transmissions. A frozen or static Ghost could indicate signal disturbance or loss.

These nonhuman, faceless, visual indications are the product of animated cycles. Animation has proved important in relation to users' perception of dynamic relations (Michotte, 1963) and their understanding of functionality (Baecker et al., 1991). The frequency of the animation cycles in and of itself may indicate the different states of the Ghost ("busy" or "waiting"), and the frequency is also synchronous with the Ghost speaking. Hereby, the user will be able to see whom he is talking to. The immediate, visible responsiveness of the Ghost's movements to the user's commands is crucial to establish a causal relationship (Michotte, 1963).

We use animated Ghosts to deal with the spatially very limited capabilities of portable displays. The animated ghosts are very like the cellular automata from Conway's "Game of Life" (Gardner, 1970). The structures possess an object identity recognizable with a very limited resolution, despite their constant transformations. For instance the "walker" 5.1 doesn't have a constant body mass, but consists purely of the movement of its "limbs."

Figure 5.1: Conway's "walker" moves from top-left to bottom-right.

The full animation cycles of the low-resolution Ghosts consist of between 4 to 20 frames of looping pattern in 8x8-pixel resolution.[1] The decision to base animated Ghosts on a very similar timeline for transitions was made to preserve a level of similarity for the overall aesthetic impression in terms of temporal regularity.

The high degree of monotony or rhythm in the figures makes deviation from the rhythm easier to detect. Easy detectability is needed to readily experience changes of state presented as changes in rhythmical movement. Starting the animation when a figure speaks and pausing the movement when the speech pauses enhance the perception of the overall, invariant rhythm in the movement. If an animation cycle contained a lengthy pause, it would be difficult to perceive whether the pause was due to speech synchronization or just a pause in the movement pattern.

[1]The timing of each animation is set to a frame rate of 15 fps., the frame distribution may be seen i Appendix A (see page **??**). If an animated Ghost needs pausing inbetween image transitions, the pause stays true to the frame rate. This approach differs from the normal capacities of frame-by-frame motion graphics, in that you would normally allow the timing of each image to be calibrated to the figure movement, rather than structured to an overall temporal constraint. In an animated gif, an image may remain unchanged from 0.1 to several seconds.

Some of the Ghosts in development during 2003-2004 were attached to named characters with established services. Others were developed for experimental purposes (see p. 111).

The animated Ghost's can be divided into four species 5.2:

- Constant body mass
- Composite body
- Physical
- Abstract

BODY

REPRESENTATION

	Constant	Composite
Physical	Physical Joe Turner Struggler Sea	Printer Jan
Abstract	L-figure	The Butler Thin Lizzy O-bar Lovers

Figure 5.2: Ten Ghost figures are introduced, based on the division of the four species: constant body mass, composite body, physical, and abstract.

Constant physical.

During the development of the DELCA project, the Ghost shown in 5.3 was named "Physical Joe." Its service was to help people remember to perform physical exercises during the day, by recommending the stairs instead of the elevator, or suggesting exercises such as stepping away from the PC to do stretching.

The figure can be described as having a constant body mass, and it actually looks physical: the mechanical effect of one side diminishing while the other expands follows a recognizable physical pattern that excludes it from being abstract. It resembes a running engine or a marching creature.

Figure 5.3: "Physical Joe" looks like it's doing exercises. It is supposed to help people remember exercise during the workday.

The "Turner" figure resembles a rotating rectangle 5.4. It gains object constancy by rotating a stable object. The impression of rotation is reinforced by the rectangular hole in the center of the object, which maintains alignment with the figure during rotation. The figure has constant body mass and may look like a rotating mechanical component or a spinning thrown object.

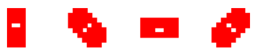


Figure 5.4: The "Turner" figure is perceived in terms of object mass. The overall shape doesn't suffer from the heavy diagonal transformation.

The "Struggler" figure starts out looking like an upward arrow and aspires to become a legged figure. This figure is basically an abstraction of the Space Invaders creature known from early computer games. It plays with the sensation of an elastic body struggling with gravity. As the ten-frame animation progresses, the stubby figure hoists its bulk into the air, revealing two thin lines. At frame 7 the two lines lose ground attachment. The figure moves too slowly to give the impression of jumping, and it is being held at the border of achievement, struggling with its own weight. The figure has constant body mass and is physical in appearance; any stout creature could conceivably act the way this figure does.

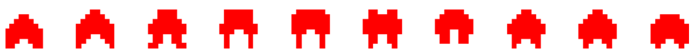


Figure 5.5: The "Struggler" struggles to heave its massive bulk into the air.

"Sea" has a constant body mass and is physical in appearance 5.6. The 12-frame animation resembles a fluid swelling in a wave-like motion. The animation has the same mirrorlike appearance as "The Butler" 5.8 and Physical Joe, a linear pattern repeating itself in symmetry. The animation touches upon a readily perceivable effect and gives the impression of borders governing the effect.

Figure 5.6: The "Sea" figure reveals that physics are involved in its animated state. The body mass is very flexible, almost fluid.

Composite physical.

Another Ghost service is "Printer Jan." It may find a nearby location for printing based on the user's location or heading. It could also handle documents at to their discretion, such as not printing a private letter until the user was present. This figure has a composite body but is definitely physical in appearance 5.7. The elements composing the figure look like a top-fed printer in action, with a piece of paper being put on top of a stack of paper in front of the printer. At low resolution, this figure transforms its divided components into a functional mass. The overall concept of the animation works around three physical elements; the visual construct of substance matter is only seen in motion. In frames 1 and 2 the paper and printer are distinguishable, but in frames 3-5 the paper stack and printer melts together. The animation is flawed in frames 3-4, as the stack suddenly decreases in height. The reason for this apparent error is to make the stack appear to grow from the paper added to it. In a looping animation it would make for a visual paper jam to add paper without removing the one previously added.

Figure 5.7: "Printer Jan" manages print queues and more.

Composite abstract.

The most important character in Ghost society is the "Butler." Its role is to guide people around and be omnipresent in an environment, which makes it very busy indeed 5.8. The figure is abstract with a composite body mass. The visual transition occurs by means of two square surfaces occluding one another. The occlusion effect forms a hole in the middle. Both surfaces visible without occlusion occurs in only a single frame. The smoothness of the 16-frame animation supports the visually complex filtering of the negative figure appearing at the center of the animation.[1]

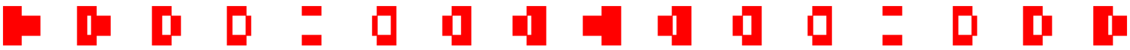


Figure 5.8: "The Butler" guides people around. He is very busy and consists of 16 frames.

The next figure is meant as a conversational agent. "Thin Lizzy" resides in the cantina or cafe and comments on the day's menu or tries to tempt patrons' sweettooth 5.9. The figure is abstract and sports a composite body. Only the first frame represents a full body. In frames 2-4 the figure is divided into four parts, each chasing the tail of the next. The constant movement supports the overall circular impression of the figure.

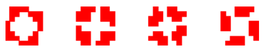


Figure 5.9: "Thin Lizzy" resides in the cantina and recommends meals. The rotation of elements ensures figure integrity.

The "O-bar" figure, which metamorphoses from two bars to a circle and back again, also plays with our sense of object constancy 5.10. This figure has two states, a two-bar state and a circle state. Both states are highly recognizable thanks to their simplicity. So the figure maintains its dual state even though five of its six frames are joined in a circular pattern. The figure is abstract and composite. The figure can be interpreted as either one figure consisting of two bars or two bars joining to form a single figure.

[1]Negative space is an old Gestalt illusion, as found in the Kanizsa triangle, which is formed by negative space.

Figure 5.10: The "O-bar" figure. The life cycle of the animation breaks up and joins two bars into a circular figure. The figure survives this transformation well in spite of the fundamental change.

The "Lovers" figure 5.11 consists of two squares (one larger than the other) rotating around each other. At 180 degrees of rotation, they seem to melt together and switch places. In the 12-frame animation, the figures are only joined in two of them. This produces the overall impression of two figures in a symbiotic system. The animation plays around with object constancy and planetary movements. It is of composite body and is rather abstract even though there are objects in the real world that could behave in this manner. This figure has not been associated with any service yet.

Figure 5.11: The "Lovers" figure displays a mutual relationship between two figures that produce the overall impression of a single figure.

Constant abstract.

The "L-figure" 5.12 uses the same overlapping technique by moving two elements as does the Butler 5.8, but instead of applying negative space, it has two elements overlapping each other, to be interpreted as a single figure. The animation has a constant body mass and is abstract. The impression the figure is supposed to give is of a limping or nondirected action. There is no recognizable feature from real life, except maybe two elements held together with some kind of goo. The figure challenges the concept of object stability by employing a dual composition or elasticity. This figure has not been associated with any service yet.

The main difference between the DELCA approach and the traditional HCI "agent" (e.g., Smartakus, Mob-i, and Rea [(Reithinger et al., 2003), (Marcus and Chen, 2002), and (Cassell et al., 1999), respectively] is the intensive use of auditory communication supplied with timely minimal visual cues. The character is

Figure 5.12: The "L-figure" employs an overlapping technique to produce motion. The resulting animation gives the impression of object constancy by its lack of rigidity of mass.

humanlike by virtue of their vocal expressions. Visual cues as to their presence are kept to a minimum (hence "Ghosts"). The Ghost metaphor acts as an immediate explanation for the characters' lack of bodily presence in the visual domain, and it sets up a range of acceptance and expectations pertaining to the service's ability to be functional despite being disembodied. The combination of context filtering with the perceptually high communicative format of voices supports the characters' presence. In this way, the ability of a seemingly omnipresent Ghost service to follow the user around in a building, jumping from display to display, hopefully becomes easier to understand.

The nonhumanlike, faceless, visual indications of entities are the product of animated cycles. These cycles are functions of their existence and capacity, rather than references to static substance. The activity state is visually communicated by cycle frequency and motion pattern. The low spatial resolution is compensated for by use of the temporal dimension. The Ghost figure may briefly escape its normal visual pattern in order to communicate urgent messages. For instance, a Ghost may need to transform into a pointer in order to point out a location. This transition of the visual appearance would traditionally be perceived as two individual signs. The constant molding of the animated figure enables it to remain conceptually present even when briefly morphing into a different visual pattern 5.13. This feature is based on our perception of object constancy in relation to motion (see p. 77).

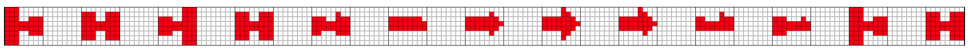


Figure 5.13: "Physical Joe" morphs into a arrow in order to point to the stairs.

I have now introduced a small gallery of ten Ghost figures. In a complex environment, there may be dozens of such services, each represented by a Ghost, but any user might only be familiar with a given subset of the Ghosts. In the next sections, I explain in more detail how the Ghost figures could inhabit an ambient intelligent environment.

5.4 The habitat of the Ghosts

Positioning technology, or location computing, is a field that encompasses several technologies used in registering the physical location of people. These technologies may use various methods to register location, such as specialized devices or tags or wireless network protocols such as active badges, Bluetooth, radio frequency identification (RFID), wireless local area network (WLAN), and optical tracking. I do not go further into the differences between the technologies, as each has its own advantages and disadvantages regarding spatial and temporal resolution and power consumption. The common factor in these technologies is the fragility of display methods to communicate information.

The system is based on positioning the wireless devices worn by the user by triangulating the client signal strength from nearby base stations 5.14. The entire area is calibrated by measuring signal strength within a grid structure with units of, for example, 3x3 meters. By comparing the user's signal strength to the calibrated information, the user's location can be deduced, but the signal strength may vary dynamically because of the changing position of physical barriers relative to the user 5.15. Position can be established with a precision of 1-2 meters.

On top of the positioning data, a logical model layer is added to the system. This type of layer contains such information as the most frequent routes taken by a user, which may make up for the lack of precision of the positioning data 5.16. The logical layer also contains structures not mirrored in the physical layout, but defined in an independent virtual environment of Ghosts and local areas, in which special rules may be applied. A local area can be conceived as an interface to services: the mere presence of a user within an area could trigger actions by the system.[1] For example, walking into the parking lot can trigger a Ghost who

[1] In this sense, the physical world could be compared with a website. Entering a specific area is like activating a link by a "mouse-over" function

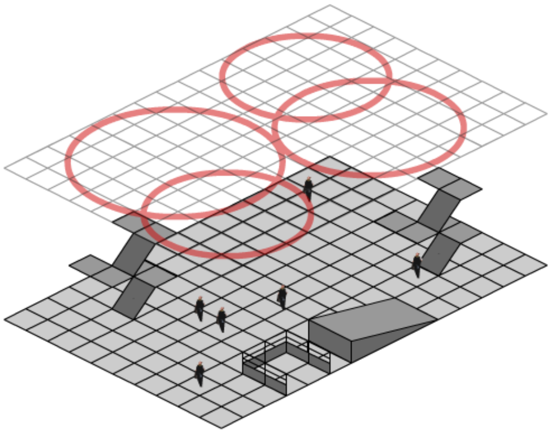


Figure 5.14: Network base station coverage. For the sake of example, the coverage is depicted as perfect circles; in reality coverage varies because of physical barriers.

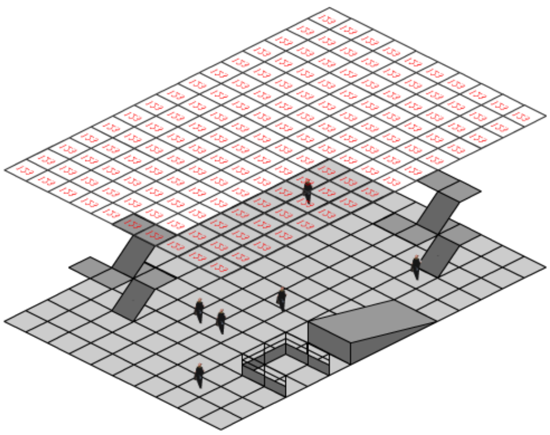


Figure 5.15: Calibration grid. Each field within the grid has a measured value used for positioning the user.

fetches the latest traffic report. In this way, the physical layout of an area can be used as an interface to digital services. Some areas could be customizable at the user level, creating a personalized virtual mapping onto the physical environment. A user could instruct his workstation to print confidential documents only when he is present in front of the printer, eliminating the risk of sensitive material lying around unattended.

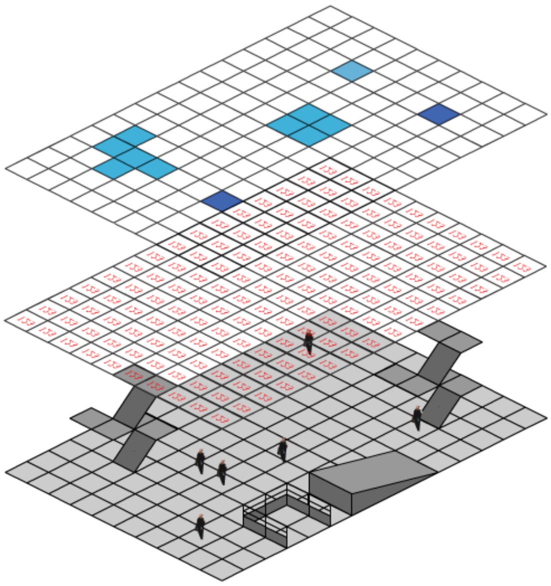


Figure 5.16: The logical layer assigns areas and services to the physical setting. The areas may be constructed independently of physical layout.

The layers' merging of physical and virtual architecture constitutes a hybrid space, which is neither physical nor virtual, but composed by experience.

Despite their obvious lack of intelligence, Ghosts may appear "undead" because of special service components. A Ghost can spice up a conversation or fill a gap in dialogue by commenting upon localized information such as the weather, thereby supporting the impression of presence 5.17.

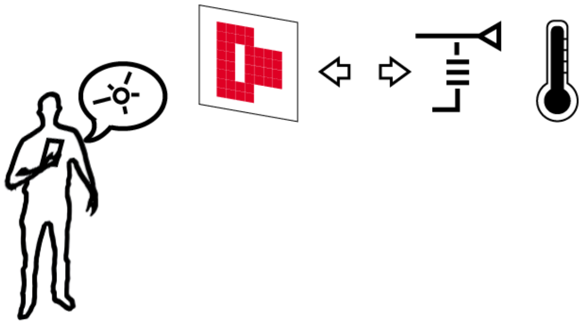


Figure 5.17: Accessing a Ghost constitutes accessing a specialized service. A Ghost may add liveliness to user conversations by interjecting sensory data or information gathered from the network. In so doing, it may reflect both virtual and physical phenomena. Illustrated here is a Ghost accessing meteorological data from sensors and using them to comment upon the weather.

5.5 Interaction with Ghosts

The services, or Ghosts, are individually recognizable agents that carry out various tasks or offer their services in various contexts within a local area. They can appear in four basic situations:

- An omniscient Ghost is available
- A localized Ghost is detected
- A specific area is entered
- Interaction with a Ghost is already in progress

The experience from the Brewster experiment, which used earcons to relate navigational cues to abstract structures (Brewster, 1998) proved inspiring for the design of the DELCA Ghosts at two levels. First, the ability to easily recognize a Ghost thanks to sound or a visual signature, or a combination of the two, is vital to the speed at which a user may interact with the interface. Also, earcons may provide the user with less dependence on the visual display in situations formerly requiring visual mode. When making selections from a complex hierarchy, earcons may help the user to group information, and in so doing narrow down the selection process by making it an option to maneuver the information hierarchically.

Second, earcons are used in the DELCA system to create redundant or secondary information for the locational and functional properties of the Ghosts and environments. A Ghost service processing a print job may emit "productive" sounds to keep the user notified of the progress, or an ambient sound may be used to create atmospheres at certain locations to induce moods in the user. In the same way earcons have been used to navigate abstract structures, sound could provide the user with auditory cues as to their position within the physical construct. This would have value not only to blind users, but also to people engaged in activities other than visually monitoring their locational information while moving around in an area. For example, while driving an automobile, the user could be given vocal directions, which are then supplemented with earcons signifying other aspects, such as service stations coming up, at which point the user may decide to utilize WLAN services.

The Hubbub method of deploying sound to communicate user presence on the messenger network (Isaacs et al., 2001) is relevant to the DELCA Ghost system, in that it enables nonfocused communication as well as represents users or agents not physically present but interacting at a social level. The ability to interact using minimal cues of identity supports the high interaction speed required for daily-life interactions. The methods for building a library of individuals connected to recognizable sound and visual features are similar to the DELCA Ghost approach of assigning visual cycles and ambient sound or voices to Ghost characters.

The display as such becomes secondary to most interactions with the system. It provides confirmation of processes in progress, notification of possible actions (i.e., when a specific service is available), a gage for the energy level of network traffic or the battery, or a measure of the potentiality of the service. At the auditory level, this could induce the voice of a very localized Ghost to address users in a feeble whisper, while a more powerful, omnipresent Ghost may speak with loud confidence. In the visual sphere, similar kinds of affective states may be visually instigated by use of varied energetic patterns or weak shaded signals. A popular Ghost may be popping around tremendously, making it difficult to reach because of demand. On the other hand, if the Ghost is suffering from a lack of computational power on the client unit (low battery or low memory), it might present itself in the auditory dimension with an almost depressed and slowed voice, while at the visual dimension, the basic cycles of the Ghost animation might be similarly slowed.

5.6 Causal relations to the Ghosts

By timing the visual occurrences with user interactions, causal relations to the dynamic environment are established. When a user asks the Ghost Butler to show him the direction to Mr. Hansen's office, wall-mounted electronic displays light up to guide him along 5.18.

Figure 5.18: Information may be displayed on small, inexpensive 8x8-LED displays placed around the environment.

> Mrs. Jones starts walking, but heads in the wrong direction. "Excuse me, Mrs. Jones. You are not going in the right direction," the Butler comments. "If you need directions, say 'yes.'" Mrs. Jones says, "help," and the animated ghost pattern appears on the display."Follow me please. I am now on the wall display." Mrs. Jones looks up and notices the same pattern on a wall-mounted mini display some meters ahead of her. She walks toward the animated figure. "That's the way to go," the Butler comments.

A major challenge of designing visual information to be shown on displays placed in the environment is controlling the circumstances in which people will look at the display. Many factors may influence visual communication, such as the angle from which people view the display, the surface at which the information is applied (in the case of projections), and, of course, the distances or scale at which information should be viewable. I argue that the visual information displayed must be scaleable, from very tiny to very large, and must be able to survive being at odd angles or on multitextured surfaces. They should still be recognizable even when projected onto surfaces like folded curtains or when parts or corners of the figure are missing because of occlusion 5.19.

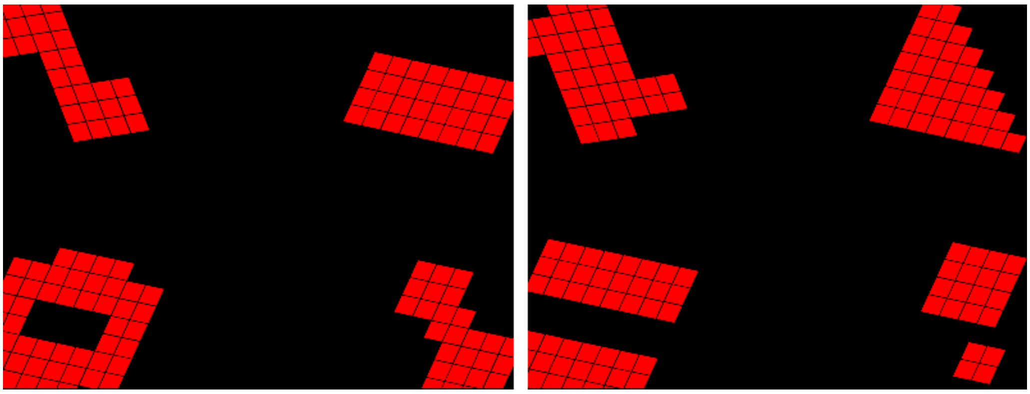


Figure 5.19: Ghost figures partly cut off and distorted by perspective. The figures are in fact "L-figure" 5.12 (top left), "Sea" 5.6 (top right), "O-bar" 5.10 (bottom left), and "Lovers" 5.11 (bottom right).

The design for presenting event relations is essentially a causal technique. As previously discussed, relations can be readily picked up by people if timings are considered (Michotte, 1963). The causal technique in public display interaction seems particularly strong for:

- Personal-display to public-display interaction
- Public-display to public-display transitions

5.6.1 Personal-display to public-display interaction

The first case of personal to public displays can be applied in guiding tasks; here, the link between the abstract-portable interface can be supplemented by external displays available in the surroundings. As seen in the parking-lot example 1.3, a relation between a portable interface (the car alarm remote control) and a display available in the surroundings (the sound of the alarm) can be combined to produce highly effective guidance information. The same basic principle could be applied in hybrid space: a portable device like a PDA or mobile phone could convey interaction with a service. The primary mode of communication could be audio, freeing the attention of the user from the visual interface of the device to the physical surroundings. In cases of ambiguity, the user could ask for a visual indication of the direction, in which case the service could light up a public display in synchronization with audio information given by the device 5.20. When we have to switch on the light at a location new to us, we employ the same basic principle. People tend to try out light switches until they experience a correlation between the flick of the switch and the light turning on at the desired location. The response of the light must be sufficiently swift for us to perceive the correlation with the switch. A common situation in which this response timing is delayed, breaking with our perception of causality, is in the cases of locations with neon lighting. Neon lights take a while to activate, and consequently many people tend to flick several switches, resulting in the loss of the perception of causality. This also manifests itself when people have lit too many lights;when they then try to switch off the unwanted lights, they often accidentally switch off the light they wanted in the first place.

The personal device can function as a remote control for public displays. Respecting the timings necessary for creating causal relations could produce swift interactions with networked environmental features such as lamps, displays, and doors. A mixture of contextual filtering and timely visual indications could lower the reliance on visual confirmation from the portable device, and make sharing public displays possible. In the parking-lot example 7, the assembly of cars could be said to constitute a public information space in which each car corresponds to a single display. The unique property that identifies when the display is responding to a user lies in the interactive property of manipulating the display - pressing the alarm's remote.

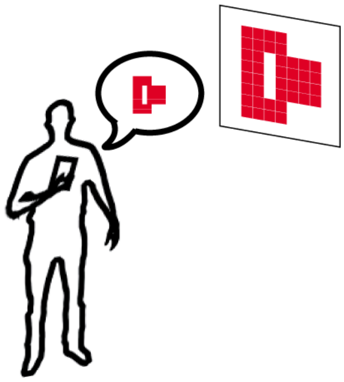


Figure 5.20: The timing of interaction or sound with that of visual cues may produce an experience of causality.

5.6.2 Public-display to public-display transitions

Moving around in physical space will change the displays most suitable for communication, as happens when any other feature of the environment changes its position relative to the user. While the changes to the physical environment inflicted by movement are evident to people in most cases, as when moving across a room makes the proximity of one chair more inviting than another, the changes to the states of a virtual service are harder to perceive.[1]

What is needed is a discrete method of making people aware of a service, as a response to their actions, even when those actions are not directly related to the service. For example, a service may change location by means of switching displays in order to follow user movement, or a service may switch displays because

[1]The confusion about displays is based on their ability to present information directed at a particular person, but with very few nonverbal cues. Imagine a meeting with several people that is about to commence, and suddenly one person tells you, "Your wife called." Normally, this information would be perceived by the intended person according to the numerous nonverbal cues in play among the other people present: looking someone in the eyes and stating something will often make that person feel that the statement was directed at them personally. Equally important, the reaction of the person receiving the information is indicative of whether that person was in fact paying attention. In the event nonverbal communication and responses are unavailable, other strategies for communication must be employed. In the case of speaker announcements, where the person speaking cannot receive response from the person intended to receive the message, the limited interaction often results in formalized syntax: "Mr. Jones is wanted in reception." While it will get the job done, the use of third person is not ideal for the user's mind-set of being at the center of the world nor is it discreet.

the display is currently occupied by another service interacting with the user.

The challenge of switching displays is how to do it in a way that is obvious to the user but still ambient. On top of this challenge, displays might be far apart from each other, requiring the service to cover large distances. Virtual services are not hampered by physical distances, but the fact that these transitions are secondary to user interactions requires a variation of the causal technique applied on personal-to-public displays. The personal-to-public displays technique applies direct transitions from one screen to the other irrespective of the distances involved. I suggest that public-to-public display transitions should perform a little differently, by involving techniques known from cartoon animation.[1] The basic principle put into use is the easing in and out of movement and the loading of movement (e.g., moving in the slightly opposite direction right before launching 5.21). These basic visual tricks make it possible for the user to pick up transitions that occur automatically. The cartoon techniques may act as cues for what is about to take place, and the user is more likely to accept the sudden emergence of the figure on another display as linked to this transition. Transitions made by mechanical bodies flying through empty space invoke the effects found in the paradoxical cases of object removal, in which movement itself is experienced without the need for visual confirmation (Michotte, 1963). The human ability to interpret object removal in this way is also found in research with objects that are perceived as disappearing behind a screen and reappearing on the other side (Gibson, 1979). I argue that the phenomenon of object removal may be accepted by people even when removal takes place in apparent thin air. In this way, displays become "magic lenses" that may reveal parts of our surroundings otherwise hidden to us (Bier et al., 1994).

5.7 Ghosts on small displays

Visually announcing the presence of a service raises the question of how to do so without occluding other information. I propose a technique named "Ghost

[1]Cartoon techniques of movement are being increasingly put into use in interfaces. One such example is the elastic effect of Windows appearing in Apple OS X or its bouncy icons.

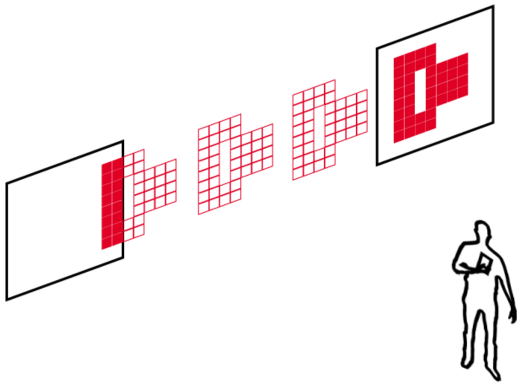


Figure 5.21: Public-to-public display transitions involve the use of cartoon techniques for discrete communication of the event. The figure in between the displays is invisible until it reaches the new display.

Wake," which is capable of displaying figures of visual significance without occluding already-existing information by means of a simple temporal transformation on parts of the interface 5.22.

The Ghost Wake technique resembles the act of moving an object behind a thin cloth, revealing the object solely through the displacement of the cloth. In this way, the Ghost lives in an invisible world behind our normal interfaces, and we may just get a glimpse of it. In our case, we make small displacements of pixels. Ghost Wake visualizes dynamic relations to the environment with a minimum of intrusiveness, without occluding the primary task. Temporal changes are immediately recognized by the user.

The Ghost Wake effect relies on the "accretion" of optical texture (Kaplan, 1969)[1] The accretion effect technique differs from traditional accretion-deletion, as the texture of the occluding surface is equal to the texture of the occluded surface. The texture defining the occluding surface is a spatially displaced version of the occluded surface. The experience of the occluding surface can only be

[1]The accretion effect occurs when a surface is partially occluded by another surface of equal texture. The experience of two surfaces can only be experienced through motion - motion either by the user or by one of the surfaces. If surface B is partially occluded by surface A, as surface A moves, an accretion-deletion effect can be experienced in one of many ways: deletion of texture from surface B as A occludes it, accretion texture of A as it occludes B, accretion of B as it reemerges from behind A's movement, and deletion of A as it moves away (Reed, 1996).

achieved by the dynamic change of texture on the occluding figure. The occluding edge transition defines the figure by temporal means. The figure itself has only the texture it occludes, and as such it is invisible in a static state (Gibson et al., 1969).

The Ghost leaves an imprint in the primary task interface that carefully moves with the Ghost. The imprint is only seen temporarily, because of the movement of the Ghost. There is no other change to either color or composition of the primary task interface.

Ghost Wake may be employed on mobile displays when a new service becomes available or when the user enters a logical area. The virtual interface becomes the "dragged cloth" that reveals invisible characters. The visual effect is sufficiently discreet that it is only noticeable when the user's attention is already directed at the interface. If a more attention-demanding effect is needed, Ghost Wake could also penetrate public displays when the user passes them or it could trigger auditory cues /The Ghost Wake effect can be experienced on the CD-ROM).

Figure 5.22: A Ghost passing by may be revealed to the user on his PDA while he is engaged in another task, the Ghost Trail causing spatio-temporal displacement of the primary-task interface. To the right, the principle is illustrated by a pattern being broken by the displacement of a circular part of it. To the left, the effect is hardly noticeable as motion is absent in the image. The effect is difficult to replicate with static media such as paper, as the displacement is minimal and is much more readily perceived dynamically thanks to our sensitivity to motion.

The Ghost leaves a temporary imprint in the primary-task interface, but it is very hard to see the actual identity of it because of the random composition of the image it displaces. The design solution for showing permanent Ghost figures was therefore to use a limited margin of the PDA to do so. Areas are signified by gray transparent areas, and services by tiny animated symbols 5.23. Both the area and the Ghost can be shown in monochrome and require little detail.

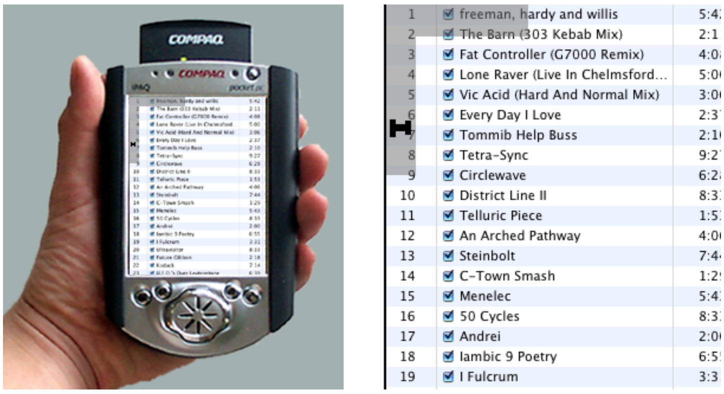


Figure 5.23: Images of an MP3 listing and an area containing a Ghost.

5.8 Summary of design proposal

A design concept for Ambient Intelligence has been presented. It aims to bridge mobile devices and ubiquitous displays, and to follow the user through the environment. The visual form consists of animations of low-resolution figures, because they are robust against distortions and because they take up little space on a mobile display. A "Ghost Wake" technique has been presented, which allows the comings and goings of Ghosts to be visible without disturbing the primary-task interface.

In the next chapter I report an experiment conducted to test parts of the design.

Chapter 6

Empirical evaluation

6.1 The experimental setup

A test was performed online by 40 volunteers at the ITU. The average age of the subjects was 32 years, with an equal number of males and females. The test program was made in Macromedia Flash.

The experiment consisted of an omnibus of smaller experiments. During the experiment users cycled through seven different tasks. The experimental session lasted approximately 20 minutes in total. Data analysis was performed by John Paulin Hansen.

The subjects were randomly divided into four groups that interacted with either:

1. cyclically animated Ghost figures in normal view 6.4 (see p. 114),
2. cyclically animated Ghost figures in distorted view 5.19 (see p. 104),
3. blinking Ghost figures in normal view 6.4 (see p. 114), or
4. blinking Ghost figures in distorted view 5.19 (see p. 104).

All subjects were given basic instructions for proceeding with the experiment. The subjects inputted their initials, age, and gender, and were given a hidden session-id number by the system for identification 6.1.

Figure 6.1: Initials, age, and gender input.

To avoid errors caused by sound levels being too high or too low, thus possibly degrading performance of the next screen, the subjects were played a synthetic voice suggesting that the volume of the speakers could now be adjusted.

To ensure that the subjects could navigate properly using the keyboard, the following screen provided a practice session, instructing subjects to press the numeric keypad keys 7, 9, 1, and 3 in response to a changing pattern on a grid figure 6.2. Subjects were required to press the key that best corresponded to the position of the red square. The key mappings corresponded to the red square's placement in the grid - top-left, press numeric key 7; bottom-right, press 3, and so on. After ten correct keystrokes, subjects would automatically proceed to the next screen.

The following screen, 6.3, informed subjects of the next task, which was to identify the figure speaking from among four possible candidates by pressing the numeric key corresponding to the figure's position - 7, 9, 1, and 3.

6.2 Experiments 1 and 2 reaction times and errors

The voice-figure tests (Experiments 1 and 2) consisted of four figures being displayed, one in each corner of the screen. One voice would be heard for each interaction, and one figure would move in synchronization with the voice. The

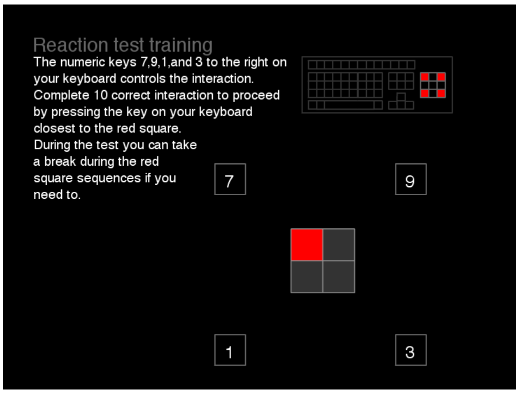


Figure 6.2: Training on key mappings.

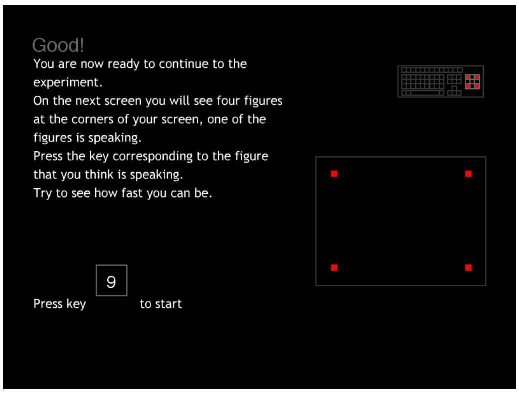


Figure 6.3: Introduction to the voice-figure tests.

figure would move (blink or animate) when the voice began, and stop the relevant motion when the sentence was finished 6.4. In every case, the sentence was: "(Annie, James, Lucy, Peter) has a message for you. Please pick it up at your convenience." The voices were generated by the synthetic voices available from Rhetoric and ATT.

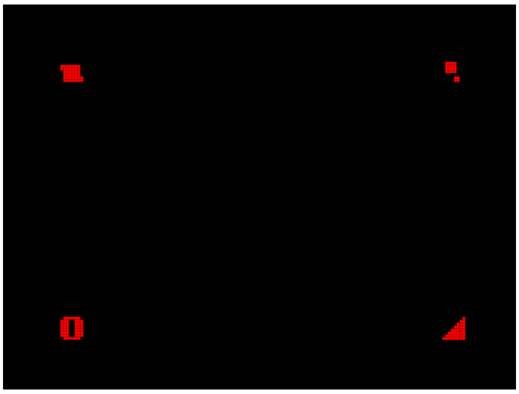


Figure 6.4: Voice-figure test. In each iteration, one of the figures would display motion that was in synchronization to its voice. This set is in normal perspective. A skewed perspective is shown in 5.19 (see page 104).

This first part of the experiment required 64 selections (4 figures x 8 iterations x 2 states). At every iteration, the figures would change places randomly.

The two states figures could speak in were as follows:

- No distractors: one figure speaking and moving (animated or blinking) while the rest of the figures remained stationary
- With distractors: one figure speaking and moving (animated or blinking) while the rest of the figures moved in sequence with random starting points

Measures were taken to preclude a figure from speaking twice in a row in the same state. However, a figure could speak successively with, and then without, distractors (and vice versa).

Data derived from experiments 1 and 2 consisted of the figure speaking, its location, whether it was correctly identified by the subject, and the time it took to identify the speaking figure. Each of the 64 repetitions was paused by the task of reacting to the red square placement already encountered in the introduction, though only one correct selection was needed. This interlude was conducted to make sure subjects were forced to break any tendencies to favor a particular corner placement.

User interaction during experiments 1 and 2 was aimed at building a relation between figure and voice by means of a test that formally measured reaction time. Experiment 1 was conducted in a complete random mix with experiment 2, the display of figures without distractors, with only one figure that moved.

Results of experiment 1 The average response time measured 1643 ms.

> A main effect from the form factor was found in an analysis of variance (ANOVA), F(1,39) = 5,59, $p < 0.05$; subjects were treated as iterations. The target Ghost was less rapidly identified under the normal perspective than it was with the skewed perspective (ms = 1945 and ms = 1342, respectively).

The distorted version resulted in the fastest response times. This may be due to the larger size of the figures being an important factor in people's ability to discover new information. Further experiments should be conducted to disclose whether size alone causes this result. Two experiments are envisioned, one in normal perspective/normal size and the other in double size/normal perspective.

The average rate of error in identifying the figure speaking was 15 percent with beginners. Even though this error rate is relatively high, one should remember that this was subjects' first encounter with Ghost figures. Obviously, longer successive blocks of experiments should be performed in order to trace the learning curve. No effect of either form (normal vs. skewed perspective) or movement (animated vs. blinking) was found in the error rate.

At this time, I cautiously conclude that the normal distortion of figures projected does not seem to add anything to the response time. The speed at which distorted figures are identified is most likely due to the exaggerated size often produced by projections. At odd angles, figures often seem larger than they are with normal perspective.

Results of experiment 2 From experiment 1 there seemed to be little effect whether figures blinked or were animated; in a situation in which only a single

Ghost was speaking, effiency was gained by blinking instead of animation. However, in experiment 2 with multiple distractors, blinking was easier to identify quickly than was animation.

> Reaction times: The grand mean of reaction was 2864 milliseconds. A main effect from the visual form factor was found F(1.39) = 11.39. p < 0.005. Target was slower to identify under the animated condition (ms = 4220) than under the blinking condition (ms = 1506). No other effects were found.
>
> Error rate: The grand mean error rate was 0.29. This was a major effect from visual form F(1.39) = 8.59. p < 0.01. The error rate under the animated condition was higher (ER = 0.43) than under the blinking condition (ER = 0.17). No other effects were found.

The introduction of distractors caused response time to almost double compared with experiment 1. This increase was mainly caused by animated figures being much harder to detect than blinking figures. In fact subjects detected the blinking figures just as quickly in this situation as they did on average in the situation without distractors.

The rate of error is equally affected by the introduction of distractors. Again this is mainly due to the animated figures.

> The grand mean error rate was 0.29. This was a major effect from visual form F(1.39) = 8.59. p < 0.01. The error rate under the animated condition was higher (ER = 0.43) than under the blinking condition (ER = 0.17).

Animated figures resulted in an error rate of 40 percent, which was the same error rate as when they were shown without distractors in experiment 1.

Animated figures have a hard time competing with blinking figures when there are several of them present in the same display. Changes in an animation's cyclic rate caused by changes in the speaker's rhythm aren't as easily detectable as changes in the pulse frequency of blinking figures.

A change in visual design could strengthen animations. One possible improvement could be to reduce the frame rate of the animated, nonspeaking figures while retaining normal speed for the speaking figures, thus increasing the signal-to-noise ratio.

6.3 Experiment 3 familiarity with voices

In experiment 3, the subjects were told to judge whether the voice they were hearing was a voice they remembered from experiments 1 and 2. While listening to the voice, subjects had the task of pressing one of three buttons: "yes," "no," and "not sure." The experiment consisted of ten voices including the four already encountered. No visual support was given.

From the results 6.3, we may conclude that people are capable of recognizing well-known synthetic voices in four out of five instances (an 80 percent rate of success) after a rather short time. There were no differences related to the four experimental groups.

Group	Correct	Wrong	Not sure
1(n=10)	84	9	6
2(n=10)	86	8	6
3(n=10)	84	12	4
4(n=10)	75	15	6

Table 6.1: Recognition of "new" and "well-known" voices. The table displays the number of voices identified as "heard before" or "new" among ten voices played. Four of the voices had already been played 64 times in total.

These results suggest that voices can be used to indicate whether a service is new to a user or if he has in fact encountered it before. The use of "well-known" voices would signify familiarity. The result also emphasizes the importance of a service retaining the same voice for a specific user.

6.4 Experiment 4 figure-voice associations

In the next experiment, the subjects' task was to identify the figure speaking from a choice of ten figures. Subjects were also informed that some of the figures would be unknown.

Subjects would then see the row of ten figures and hear a voice. They were then required to click the figure speaking. In contrast to experiments 1 and 2 6.5, there was no synchronicity between figure movement and voice. The task was iterated 30 times (10 figures x 3). At each iteration, the figures would change places at random.

> The grand mean error rate was as high as 0.84. At chance level, subjects would have an error rate of 0.9, so there is no indication of retention of a name-figure relation after the 64 exposures to the Ghosts.

No significant effect from form or motion was found. This was a rather disappointing result. I had expected that subjects would at least be able to point to a figure they were familiar with from experiments 1 and 2 when they heard one of the four voices from those experiments. The explanation is most likely that there is no natural coupling between a humanlike synthetic voice and an abstract Ghost figure. If the figures had been faces, perhaps, there might have been information supporting the formation of a voice-figure link, perhaps in terms of the gender, age, or ethnicity of the voice sound. Voice-sound information is not of any use for our Ghost figures in their present form, but some visual coding schemes may be possible: older voices could be related to figures with subtle movements, for example, and younger with eager movements.

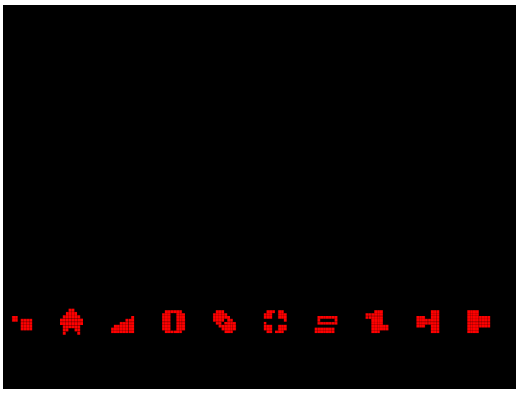


Figure 6.5: The "Who is speaking?" task.

6.5 Experiment 5 recall of figure names

People do have a fairly good memory of whether they have heard a particular voice or not, as per experiment 3. But do they actually listen to what is being said? In the introduction to the experiment, all the voices began by presenting their name and then reporting that a message was available (”Lucy has a message for you.”). In experiment 5, subjects were asked to recall as many of the Ghost names as possible and type the names of them. Results can be seen in table 6.5.

Ghost name	Remembered by	Well known/New
Mary	20	Well known
James	17	Well known
Lucy	15	Well known
Peter	15	Well known
Polly	10	New
Bill	7	New
George	7	New
Simon	7	New
Sarah	4	New
Annie	1	New

Table 6.2: Free recall of Ghost names. Also see Appendix 1.

”Polly” was among the five highest scoring, even though she was a ”new” voice, heard only four times. The other four figures in the top five were all ”well known” from experiments 1 and 2. The voice most subjects could remember the name of was only remembered by half (20) of the subjects. This indicates, like experiment 4, that no real attachment to the figures was formed during the experiments. A major reason, I believe, is that the figures all said the same sentence, and that the meaning of the sentence was in fact of no relevance to the user. Attachments might well be formed to individual figures with unique messages that relate to the user’s current task or context. Only in an in-field experiment would we be able to tell if this would happen then.

There were no significant variations between the groups.

6.6 Experiment 6 subjective evaluation

The next experimental task was to evaluate all ten figures on a seven-point scale, ranking them from "elegant" to "not elegant at all" 6.6, and on another seven-point scale, ranking them from "recognizable" to "not recognizable at all."

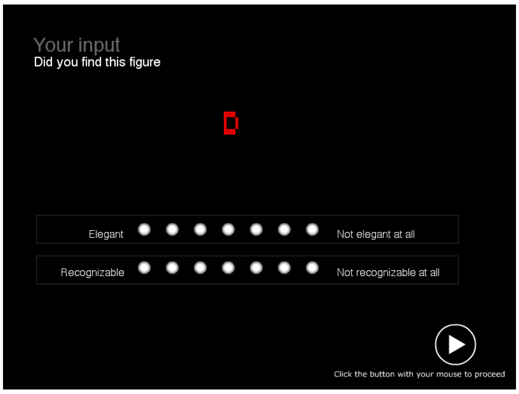


Figure 6.6: Questionnaire about the figures' appearance.

- A. Ghost figures, elegance:
  Grand mean of elegance rating was 3.80, s.d. = 1.79. The subjective ratings were analyzed by a 3 factor ANOVA, with Ghost figures (10 different), perspective (normal or skewed), and visual form (animated or blinking) as independent variables. There was a major effect from Ghost figure $F(9.385) = 4{,}41$, $p < 0.001$.

polly	3.177
james	3.202
mary	3,.408
lucy	3.753
bill	3.78
sarah	4.049
george	4.078
peter	4.459
annie	4.715

Table 6.3: Ghost figures' elegance ratings.

There was also a major effect of perspective F(1,385)= 4,63; $p < 0.05$. The normal perspective was rated better (elegance = 3.6) than the skewed perspective (elegance = 3.9).

- B. Ghost figures, recognizability:
  Grand mean was 3.18, s.d. = 1.85. There was a major effect from figures on the subjective rating of their recognizability, F(9.385) = 3.11. $p < 0.01$.

george	2.589
james	2.61
polly	2.662
simon	2.892
lucy	2.994
annie	3.277
mary	3.303
bill	3.389
peter	3.944
sarah	3.969

Table 6.4: Ghost figures' recognizability ratings.

6.7 Experiment 7 blinking versus animation

The final task consisted of selecting the figure the subjects preferred from the two shown. Eventually, all ten figures would be shown one at a time in pairs: one version blinking and one animated 6.7. The blinking and animated versions were placed randomly to the left or the right.

The forced choice between an animated and a blinking Ghost resulted in animated figures being preferred in an average of 8.2 out of 10 occurrences.

> A factor-2 ANOVA with visual form exposure (animated or blinking) and form (normal or skewed) as independent variables demonstrated a major effect from movement, F(1,39)= 10,17; $p < 0,005$. The groups who had seen the Ghost animated in experiments 1 and 2 had a preference for the animated version in 9.2 cases, whereas the groups

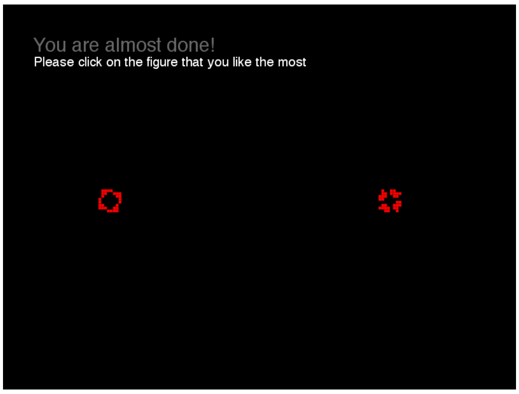


Figure 6.7: Subjects were asked to select whether they preferred the blinking or the animated version of the ten figures (the difference cannot be seen in this image).

who had seen the Ghost blinking, had a preference for the blinking version in 7.2 cases.

6.8 General discussion of results

The more subjects had been exposed to animated Ghosts, the better they preferred them to their blinking counterparts, but even the subjects exposed to blinking preferred animation.

The overwhelming preference of animation is surprisingly significant in light of the problems related to detecting animated figures with distractors. So even if the efficiency of animation is lower than the efficiency of blinking, user satisfaction is highest for animated figures. From this I conclude that animations should be used if the cost of errors or slower reaction time is low. This will typically be the case for the kind of daily activities that we envision for the Ghosts to assisting with. But the animations definitely should not be used for messages such as informing a doctor about an emergency call. In that case blinking should be employed.

Even-though animation may not be criticized for being less effective than regular symbols (Narayanan and Hegarty, 2000). Backer et al. (Baecker et al., 1991) argues that animated low resolution (20x22 pixel) icons may effectively convey the purpose of a tool. They tested animated painting icons and found that animation was productive as long as the function of the icon was not abstract. To novices animated icons better explained the functionality of a tool than static icons. With regards to complex concepts Backer et al. propose that a different style of animation may be needed when dealing with abstract concepts. The majority of situations where the use of the Ghost services are likely to occur may not allow focused interaction with computers at all. Rather than being an elaborate system designed for a complex and specialized purpose, we view the ubiquitous system as support for a series of short every day interactions together forming a highly complex system. As most of these simple interactions are only infrequently encountered by one time uses. A library of specialized icons defining each service would not work. A large number of services would need encoding by individual symbols. Infrequent users would not be able to remember the meaning of such a symbol library.

Ideally, the experiment should take place using both PDAs and external displays instead of a PC monitor. In a real-world setting, it would be possible to create natural figure distortions from the angle of projection. With spatial distribution of displays people are unlikely to be directly in front of the display. Therefore robustness to perspective distortion is a major design challenge to ubiquitous displays. We learned from the experiment that the low resolution figures where quite usable in that setting. However we also learned that the figure-voice linking needs a lot of work.

Another criticism of the experiment is the random linking of figures with voices. It would have been possible to match the voice and figure better, even in this stylistic form. In fact, one of the interesting comments included with the subjects' input was a critique of the speed of the animations for figures that were synchronized to speech. Judging from the result, even though the subject did experience a connection between voice and the figure's movement, the subject argued that the speed of the movement was too fast with respect to the voice. This was very surprising. Previous research in point light experiments demonstrated that people are sensitive to the motion patterns in moving lights, and that this visual organization may convey precise information about the person wearing the lights (Johansson et al., 1994) (Runeson, 1983). However, the relationship between the speed or rhythm of abstract animation and voice remains mostly unexplored. Future experiments should explore this relationship, as it presents an interesting aspect to the perception of figure-voice coherency. Further studies on effective mappings could draw on Runesson's and Johansson's group's research on the effects of abstracted physical form movements (Runeson, 1983) (Johansson et al., 1994) to find how these could be used in low-resolution animations. The right combination of voice characteristics with visual form could yield a stronger impression of figures in terms of gender, age, ethnicity, temper, weight, and so on.

A more clearly founded association between figure and voice would create a more vivid impression, that in turn would strengthen the individual character. My experiment failed to demonstrate memory of the figures identity when there was no meaning in the content of the interactions. But future work should not be discouraged by this finding. Since, as we have argued, it is the meaning of the situation that will be remembered. Association of figures to meaning may be established through repeated use, but they need not to be: you will be served by interacting with a bus-conductor even if you do not recognize him, and you should be served well by a Ghost, just on the basis of its verbal introductions.
One of the issues that will need to be explored further is the degree of visual transformation a figure can survive through changes of its pattern signature and while maintaining figure constancy. Furthermore by higher frame rates a more lively impression could be achieved through more accurate timings of voice and movements . The contrast between movements of figures speaking and figures simply moving should be increased. Figure size seems to have an effect on detection speed (cf. the difference between the small, normal figures and the large, skewed ones in experiments 1 and 2). If possible, figure size should be as large as circumstances permit.

Chapter 7

Toward nonencoded representation

Matching the speed at which technology can deliver information with the speed at which humans are capable of receiving and reacting to it is one of the fundamental challenges of technology. The advent of dynamic and contextual aspects of information displays in hybrid spaces has shown that the conventional ways of delivering information are lacking. This communicational shortcoming may have the consequence of producing a widening gap between human decision-making and technological filtering due to the limited communicational properties of the tools and methods in use. If we become more more reliant on decreasing the amount of information displayed to us, our ability to interact becomes hampered.

The strategy of transforming conventional symbols to new media has turned up a few problems. The "My Computer" symbol is a symptom of the fundamental problem of assigning real-world strategies to computing by the use of metaphors 7.1. Opening the "My computer" folder reveals the paradox of the computer being a symbol that contains itself.

When looking at the Desktop of the computer, how can it then contain itself? And why is it necessary to represent the computer with a symbol?

This puzzle is one of the fundamental problems of encoding, as an encoded reference must necessarily refer to an encoding. An infinite regress thus occurs, in which the content of the referred information can never be brought forward. The use of a metaphor that refers not only to the physical entity of the computer but also to the virtual information contained within it has instilled a homunculus-like logic to the information flow between user and system. The user is already likely to be sitting at a physical desk, with a physical computer also present - so why the need for mimicking the physical setting?

Figure 7.1: Screenshot of the Microsoft Windows "My Computer" icon.

Most tasks and elements portrayed visually on the computer screen are done so through the use of metaphor. Metaphors are constructed by convention to refer to information content. In a "new" medium like the computer, most metaphors derive from traditional or already-known technologies and situations 7.2. The result is a widespread adaptation of symbols and tools that fit computational tasks only to a certain extent because they derive from other media. This is the case especially for the desktop metaphor, which features the "the paint-bucket," "recycle bin," and "My computer" (icons for filling an area with color, deleting files, and the computer itself as structured by the operating system, respectively).

As interaction with the surroundings were never intended in the design, the need arose for mimicking situations as they pertained to computer usage. The desktop is indeed a true representation of early PC usage; however, the composition of that type of interface doesn't allow for a changing environment, as the reliance on copies of real-world entities would require countless symbols.

I argue that the less-than-logical constructs of the WIMP interface are only workble thanks to isolation from the real world - the need for replication of real-world features is a symptom of this isolation, not the cure. In hybrid space, the computer disappears, not as a physical object but as a window to virtual events. Virtual events inhabit physical space and blend in with natural actions instead of being obtainable only by the use of encoding.

The predominantly encoded presentation of digital information sets narrow limits to the agility with which we may interact with that information, as the

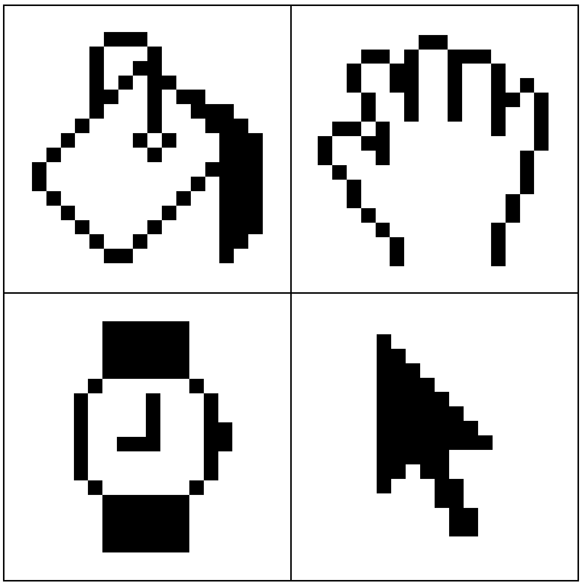


Figure 7.2: Screenshot of some of Susan Kare's classic icons. http://kare.com

encoding doesn't facilitate our embodied perceptual mechanisms, but requires translation or decoding. One may argue that by relying so heavily on encoding, we have achieved what Rodney Brooks calls the sense-think-act bottleneck (Brooks, 1991); although it originally referred to a robot vs. real-world scenario, I compare it here to the human vs. virtual-world scenario. However, in contrast to the robot population, humans have been thoroughly tested and fine-tuned by evolution - we just need to take advantage of it. In order to interact with dynamic environments, we need the creation of an interface that does not require translation or focused attention. Through our current reliance on encoding in interfaces, we are in danger of turning hybrid space into an environment unsuitable for us. As encoding requires decoding, we would constantly need to resort to a secondary system to understand and process events.

Understanding the difference between a user who is literally in a physical environment and one who is visiting the metaphorical, constructed interface of virtual reality represents a quantum leap toward an interactivist understanding of an active system.

The reentry from encoded HCI to that of the familiar environment of physical objects reinstates the ”anchors” of invariant potentialities:

> From an interactive standpoint, physical objects are epistemologically constituted as patterns of potential coordination among various manipulations and visual scans (Cambell and Bickhard, 1986)(p. 38).

The familiar world we live in is thus differentiated from that of traditional HCI, in that in HCI the array of possible manipulations lies within conventions based on encodings.

Icons are not well suited for conveying information that changes value. Icons are basically static symbols; encodings present fundamental challenges for their use. For the user primarily, the problem is one of managing the correspondences between the encodings and the elements they represent. The fundamental problem of encodings lies in the presumption of their nature as basic forms of representation:

> Encodings are representations by virtue of defined correspondences with representations. The general approach assumes that representation can be understood in terms of correspondence (of some sort) with that which is to be represented. The problem, as with foundational encodings, is to model how any correspondence could yield representational content of what is to be represented - of what is on the other end of the presumed encoding correspondence (Bickhard, 1996).

The visual cues allowing for the peripheral awareness of a Ghost service comprise animated cycles that express the different states of the service, thereby escaping some of the problems related to symbolic representation of dynamic information. For the most part, the problem of representing changes to static symbols over time will have to be deduced by the user instead of experienced. If a symbol remains static even though what it represents has changed, that changes correspondence. The different states a service undergoes could be based on many things, including activity level, availability, and response to interaction.

The conventional encoding of a value into a symbol offers a challenge in terms of conveying the possible changes of the value and in the fundamental encoding problem itself, that of knowing what a representation refers to. By allowing the services to be engaged in transformations based on state, the design aims at delivering visual cues to interaction rather than using anthropomorphic symbols.

In a superficial way, motion shares some of the characteristics as the interactive process of "internal bookkeeping," as described by Mark Bickhard and D. M. Richie (Bickhard and Richie, 1983), in that it doesn't contain encoding and yet is capable of signifying states. The power of the pixel building block of low-resolution graphics is its reliance on other pixels in order to form meaning, rather than individually containing or referring to it.

> The real power of the pixel comes from its modular nature, in that a pixel can be part of anything, from text to lines to photographs. Pixels are pixels is as true as bits are bits (p. 107) (Negroponte, 1995).

I believe that an endeavor to create a visual language that encompasses all the dimensions of mixed virtual and physical space in a single device would be futile. This belief is based on an understanding of the limits of symbols to represent dynamic relations, as we couldn't possibly create a referencing symbol for every conceivable combination of events or entities (Bickhard, 1996). Trying to represent dynamic environments through an interface that relies on encoding would run us into serious trouble. Dynamic environments contain a multitude of elements, relationships between those elements, and individual states of the elements that in turn continuously change those relationships. Representing the number of possible events through encoding would require an endless number of symbols to refer to those states. Not only would it be necessary for the user to recognize these states, but the user would also have to recognize these states by their referring encoding. The problem revolves around the fact that the interface is now directly linked to the world; as the world is open-ended, an effort to interact with it through static mapping is futile. This problem is a "new" philosophical problem called "The Frame Problem" (Bickhard).

We live in a world in which we do not know all the relations between all things at all times; rather, we generate these ourselves, through our ongoing interaction with the world. This so-called Frame Problem derives from AI research like robotics. Brooks argued that a design not relying on representation is not only possible but necessary in order to deal with dynamic environments:

> When we examine very simple level intelligence we find that explicit representations and models of the world simply get in the way. It turns out to be better to use the world as its own model (Brooks, 1991).

I argue that the discrepancy between the virtual and physical worlds invokes the Frame Problem as a major practical issue in the design of hybrid Spaces. Hybrid space changes the interface from an isolated environment consisting of a limited number of predetermined relationships to an interactive environment generating relationships between an unforeseeable number of events and elements.

The visual cues allowing for the peripheral awareness of a Ghost service present problems related to the symbolic representation of dynamic information. The encoding of a value into a symbol presents a challenge in terms of conveying the possible changes of the value and in the fundamental encoding problem itself - that of knowing what a representation refers to

As an alternative to basing design on encodings, I propose a visual language based on occurring transactions to form representations based on event perception or causality.

The ideal design space would feature the combination of location computing, portable displays, public displays, and new ways of interaction through invisible services 7.3. Location computing, in this case positioning, is made possible by the use of a wireless network. Portable displays can include mobile phones, PDAs, small laptops, or watches; public displays can be small LED displays, tiny tickers, huge low-resolution displays, LCD projections, or programmable scene lights.

In hybrid space, where virtual and physical information are combined, the main challenge lies not in the lack of screen real estate, but in the ability to convey the contextual properties of a changing environment and still enable non-contextual tasks. Location may play an important role in filtering information; however, the display of proximal relations presents the fundamental problem of referencing displays to people and events. As mentioned earlier, Bertelsen and Nielsen designed a virtual interface for a pump to be triggered by the features

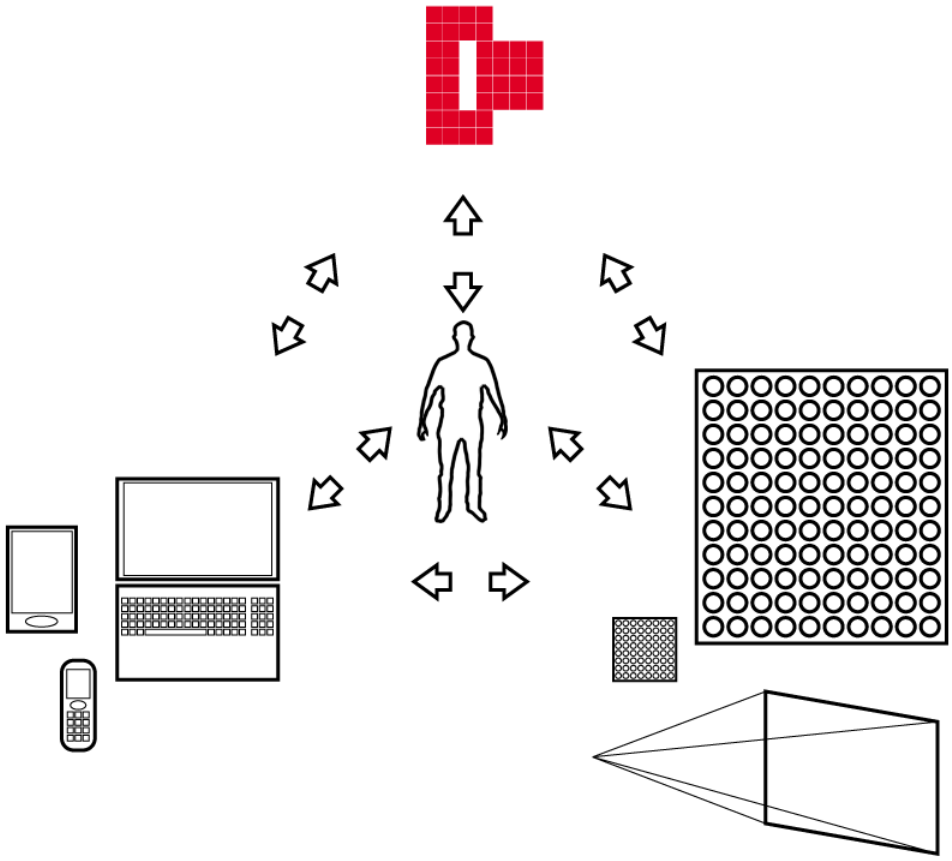


Figure 7.3: The visual design space combining portable displays with wireless positioning and public displays calls for new ways to interact. Bottom-left: PDA, mobile phone, small laptop. Bottom-right: small LED displays, projections, large low-resolution screens. Center: the human user. Top: a virtual service.

of the physical pump (Bertelsen and Nielsen, 2000). Even though the triggering was caused by the manual reading of bar codes, the lack of formal referencing in the interface could still cause confusion, as in the case of multiple pumps.

I propose that "doubt" is an invariant feature of real-world interactions, and that display systems should be designed to convey the relational properties between sender and receiver intuitively. To create and maintain a person-display connection without relying on an encoded system that maps that relationship is a major visualization task. But how is it possible to know which service we are interacting with if a formal secondary system isn't used to keep track of events?

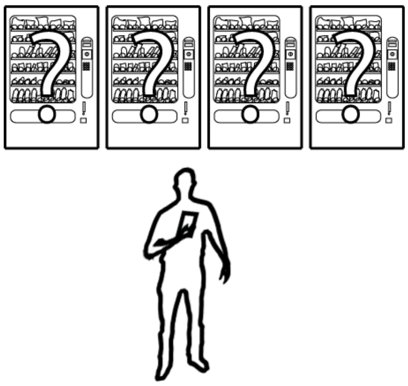


Figure 7.4: Multiple elements challenge the perception of proximal relations.

The DELCA use of external displays is rooted in the user triggering the transformation, either by continuing interaction or by direct manipulation. This involvement on the user's part creates responsiveness within the system environment that goes beyond automated occurrences and enables deictic communication, thanks to the lack of need for formal relations between sender and receiver. The deictic dimension of communication in this sense relies on the dynamically generated causal relationships between the system and the user, such as when the user presses a light switch, and the the light goes on even if the wrong light switch was pressed. The relationship is present through the negation of responsiveness. Based on temporal timings, a symbol, through the causal linking to the environment, can adopt several meanings on account of the different annotations the user might derive from its way of functioning. "A" light switch might become "a wrong" light switch. Remove or change the causal relation through change in response time, and the user no longer has a notion of what the symbol represents.

The design seeks to address a variety of services we may encounter in hybrid space, mainly services that have a relation to our immediate surroundings. Displays may take many forms: we might have to negotiate the tiny display on our digital watch 7.6, the larger phone screen, the PDA screen, monitors, or the projected display onto uneven surfaces 7.5.

Figure 7.5: Ghost projected onto the floor distorted by the odd angle.

Figure 7.6: Photo manipulation depicting a Ghost on both watch and distant screen.

Chapter 8

Conclusions

The visual design for DELCA Ghosts contains outlines for a new kind of calm visualization suited for hybrid spaces. Striving for a balance between calmness and visual significance seems to be the major design goal.

Animations can be used as powerful means to attract attention in exterior surroundings, even on low-resolution displays. They may also be used on mobile displays without occupying too much screen real estate. When the primary interaction mode is verbal, abstract animations may be a feasible alternative to face-like figures. The general idea of how to visualize personalized services without falling victim to the communicational pitfalls of facial representations seems promising.

We shouldn't continue to strive for virtual information spaces, when we need hybrid spaces, where virtual and physical space can coexist. Hybrid space makes it possible to encompass both the affordances of the physical world and the multitude of invisible structures generated by digital information. The visual design of invisible services should enable them to remain ambient when the user's intent is aimed elsewhere, yet be perceptually evident when offering services requested by the user.

When these services are designed visually to preattentive processes, people may experience services contextual to their actions and not be bothered by the rate of error at which the system guesses wrong about their intentions. Ambience becomes a functional feature when we can display information in the periphery of our attention; so long as we have no need for the information, it goes unnoticed, but when we need a service it snaps to attention. The transformation from awareness to attention cannot be achieved by visual designs based on focused attention, nor can it be achieved with ambiguous ambient cues, as these are still dependent enough on recognition to be unambiguous.

Visual design should incorporate information that results from its optical appearance rather than predetermined meaning based on recognition. A Ghost should be as discreet as a water faucet, and should be allowed to drip only occasionally, to make the user ambiently aware of its existence.

It is in the combination of different modalities that we may achieve swift transition between awareness and attention, whereas if we depend solely on one or the other, the built-in pitfalls of each modality will prove to be hindrances. Sound is effective in terms of both the informational richness of speech and the ambience of abstract or natural sounds, but it becomes inefficient if the auditory domain is already in heavy use. Shouting information at a loud rock concert will probably achieve nothing; visual cues would be far more effective. Visual information is equally strong and detailed within a perceptual threshold, after which blinking lights or signs of warning achieve little.

Skillfully balancing visual, auditory, and location cues is essential for maintaining an effectiveness and rooting in the hybrid space of physical and virtual information. However, the sensory apparatus is very adept in picking up these cues, especially when combined with timely appearances of information.

Traditional interface design and the heavy use of signs are very effective when people are converging. However, when the situation calls for quick action or multiple tasks, people are most effective in utilizing sensory information in combination with reasoning. This is especially the case when dealing with the invisible world of digital information.

The affordances between people and the environment can be used in the contextual situations in which these affordances are experienced.

Ambient systems are omnipresent. The interfaces of such systems should be designed accordingly.

Appendix A

Animated Ghost figures

The animated Ghost figures are shown here with time to frame. A blank frame signifies that the previous frame remains unchanged.

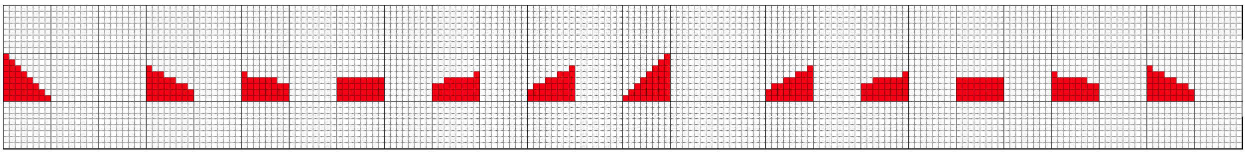


Figure A.1: "Sea" shown with frame intervals.

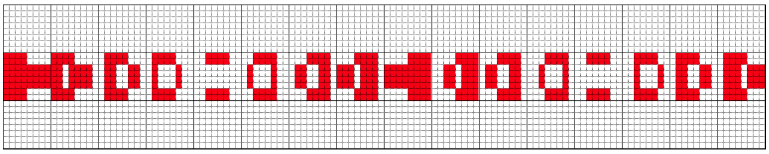


Figure A.2: The "Butler" shown with frame intervals.

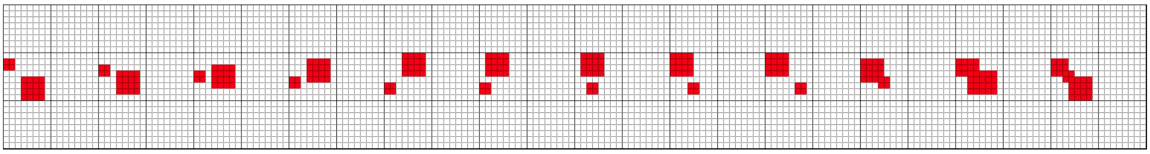


Figure A.3: "Lovers" shown with frame intervals.

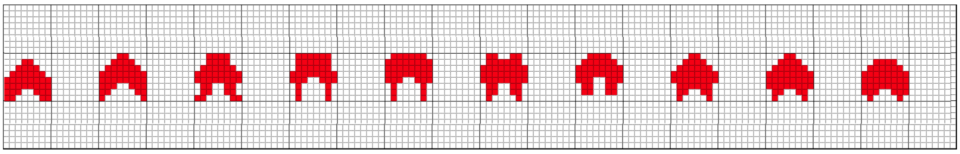


Figure A.4: "Struggler" shown with frame intervals

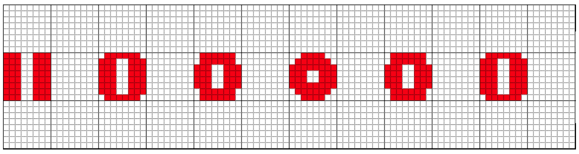


Figure A.5: "O-bar" shown with frame intervals.

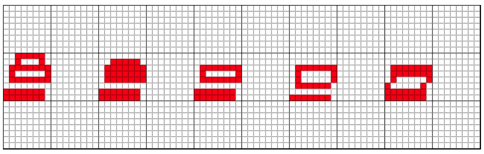


Figure A.6: "Printer Jan" shown with frame intervals.

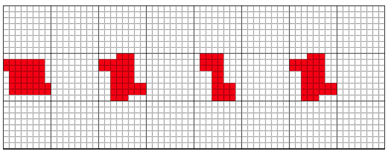


Figure A.7: "L-figure" shown with frame intervals.

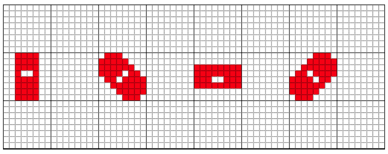


Figure A.8: "Turner" shown with frame intervals.

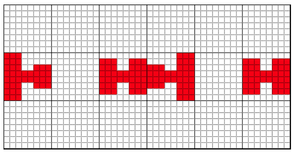


Figure A.9: "Physical Joe" shown with frame intervals.

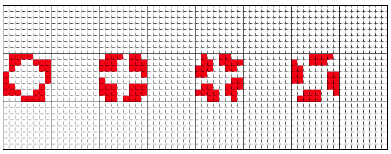


Figure A.10: "Thin Lizzy" shown with frame intervals.

Appendix B

CD-ROM

The CD-ROM contains figures and experiments mentioned in the text.

Bibliography

Gregory D. Abowd, Christopher G. Atkeson, Jason Hong, Sue Long, Rob Kooper, and Mike Pinkerton. Cyberguide: A mobile context-aware tour guide special issue: mobile computing and networking: selected papers from mobicom '96. *Wireless Networks*, 3(5):421–433, 1997. 2.3

Ronald Baecker, Ian Small, and Richard Mander. Bringing icons to life. In *Conference on Human Factors in Computing Systems (CHI '91)*, Proceedings of CHI'91, pages 1–6, 1991. 5.3, 6.8

Michel Banâtre, Paul Couderc, Jean-Marc Menaud, and Frédéric Weis. Proximate interaction of wireless appliances. 2.2

Lyn Bartram. Enhancing visualizations with motion. In *Infovis 98*, Proceedings of the IEEE Symposium on Information Visualization, pages 13–16, 1998. 4.3

Lyn Bartram and Colin Ware. Filtering and brushing with motion. *Information Visualization*, 1(1):66–79, 2002. ISSN 1473-8716. 4.3

Steve Benford, Rob Anastasi, Martin Flintham, Adam Drozd, Andy Crabtree, Chris Greenhalgh, Nick Tandavanitj, Matt Adams, and Ju Jow-Farr. Coping with uncertainty in a location-based game. *Pervasive Computing*, pages 34–41, 2003. 2.3, 2.4

Olav W. Bertelsen and Christina Nielsen. Augmented reality as a design tool for mobile interfaces. In *Symposium on Designing Interactive Systems*, Proceedings of the conference on Designing interactive systems: processes, practices, methods, and techniques, pages 185 – 192, New York City, New York, United States, 2000. 2.1, 7

Mark H. Bickhard. Why children don't have to solve the frame problems. 7

Mark H. Bickhard and D. M. Richie. *On the Nature of representation: A Case study of James Gibson's Theory of Perception.*, volume New York:Praeger. 1983. 1, 7

M.H. Bickhard. Troubles with computationalism. In W. O'Donohue and R. F. Kitchener, editors, *The Philosophy of Psychology*. Sage, 1996. 7

Eric A. Bier, Maureen C. Stone, Ken Fishkin, William Buxton, and Thomas Baudel. A taxonomy of see-trough tools. In *CHI'94 Human Factors in Computing Systems*, 1994. 3.2, 5.6.2

M. Billinghurst and Thad E. Starner. Wearable devices: new ways to manage information. *Computer*, 32(1):57 –64, 1999. 3, 3.0.3, 3.0.3, 3.5

Ole Bouman. The architectural potential of a digital world. URL http://www.c-i-d.dk/rtf/KDBoumann.rtf. 1.1

Susan Brennan, Brenada Laurel, and Ben Shneiderman. Antropomorphism: From eliza to terminator 2. In *CHI'92*, 1992. 3.0.2

Stephen Brewster. Using nonspeech sounds to provide navigation cues. *ACM Transactions on Human-Computer Interaction*, 5(3):224–259, 1998. 3.1, 3.1, 3.1, 5.5

Stephen Brewster, Joanna Lumsden, Marek Bell, Malcolm Hall, and Stuart Tasker. Multimodal 'eyes-free' interaction techniques for wearable devices. *CHI Letters*, 5(1):473–480, 2003. 3.1, 3.1

Rodney A. Brooks. Intelligence without representation*. *Artificial Intelligence 47*, pages 139–159, 1991. 7

Jenna Burrell, Paul Treadwell, and Geri K. Gay. Designing for context: usability in a ubiquitous environment. In *ACM Conference on Universal Usability*, Proceedings on the 2000 conference on Universal Usability, pages 80–84. ACM Press New York, NY, USA, 2000. 2.1, 5.2

M.P. Bussemakers and A. de Haan. When it sounds like a duck and it looks like a dog... auditory icons vs. earcons in multimedia environments. In *ICAD 2000*, Proceedings ICAD 2000, pages 184–189, Atlanta: International Community for Auditory Display, 2000. 3.1

R.L. Cambell and M.H. Bickhard. *Knowing Levels and Developmental Stages*. Contributions to Human Development. Karger, 1986. 7

J. Cassell, T. Bickmore, M. Billinghurst, L. Campbell, K. Chang, H. Vilhjalmsson, and H. Yan. Embodiment in conversational interfaces: Rea. In *CHI 99*, 1999. 3.0.2, 5.3

Franklin C. Crow. The aliasing problem in computer-generated shaded images. *Commun. ACM*, 20(11):799–805, 1977. ISSN 0001-0782. 4.1.1, 4.1.1

A. Dey. Understanding and using context. *Personal and Ubiquitous Computing*, 5(1):4–7, 2001. 2.2

Judith Donath. Mediated faces. *Lecture Notes in Computer Science*, 2117:373, 2001. 3.0.2

Paul Dourish. Seeking a foundation for context-aware computing. *Interaction*, 16(2-4), 2001. 2.2

Paul Dourish. What we talk about when we talk about context. *Personal and Ubiquitous Computing*, 8(1):19–30, 2004. 2.2

Kathryn Elliot and Saul Greenberg. Building flexible displays for awareness and interaction. In *UBICOMP2004: Workshop on Ubiquitous Display Environments*, Nottingham, UK, 2004. 3.3, 3.3.2, 3.15, 3.16, B

Thomas Erickson. Some problems with the notion of context-aware computing. *Communications of the ACM*, 45(2):102–104, 2002. 2.1

Petra Fagerberg, Fredrik Espinoza, and Per Persson. What is a place? allowing user to name and define places. In *CHI 2003: New Horizons*, Understanding the context of use, pages 828–829. ACM, 2003. 2.3, 2.4

George W. Fitzmaurice. Situated information spaces and spatially aware palmtop. *Communication For the ACM*, 36(7):39–49, 1993. 2.1, 2.2

Martin Gardner. The fantastic combinations of john conway's new solitaire game "life". *Scientific American*, 223:120–123, 1970. 4.3, 5.3

H. Gellersen, A. Schmidt, and M. Beigl. Adding some smartness to devices and everyday things. In *IEEE Workshop on Mobile Computing Systems and Applications*, Proceedings of IEEE Workshop on Mobile Computing Systems and Applications 2000 (WMCSA'00). IEEE Press, 2000. 2.1

James J Gibson. *The Ecological Approach to Visual Perception*. 1979. 1.1, 1, 4.2, 4.3, 5.6.2

J.J. Gibson, G. Kaplan, H. Reynolds, and K. Wheeler. The transition form visible to invisible. In E. Reed and R. Jones, editors, *Reasons for realism: The selected essays of James J. Gibson*. Hillsdale, 1969. 5.7

Kelly Bowman Heaton. *Physical Pixels*. Master of science in media arts and sciences, MIT, 2000. 4.2

Lars Erik Holmquist and Tobias Skog. Informative art: information visualization in everyday environments. In *Proceedings of the 1st international conference on Computer graphics and interactive techniques in Austalasia and South East Asia*, Melbourne, Australia, 2003. 3.3.1, 3.14, 3.3.1, B

Drugald R. Hutchings, Mary Czerwinski, Brian Meyers, and John Stasko. Exploring the use and affordances of multiple display environments. In *UBICOMP2004: Workshop on Ubiquitous Display Environments*, Nottingham, UK, 2004. 3.3

Infineon. Infineon and vorwerk join forces to weave the future. URL http://www.infineon.com/cgi/ecrm.dll/jsp/showfrontend.do?lang=EN&content_type=NEWS&content_oid=84620&news_nav_oid=-9979. 3.4, 3.20, B

Ellen Isaacs, Alan Walendowski, and Dipti Ranganthan. Hubbub: A wireless instant messenger that uses earcons for awareness and for "sound instant messages". In *CHI 2001*, pages 3–4, 2001. 3.1, 1, 5.5

Ellen Isaacs, Alan Walendowski, and Dipti Ranganthan. Hubbub: A sound-enhanced mobile instant messenger that supports awareness and opportunistic interactions. *CHI Letters*, 4(1):179–186, 2002. 3.1

Hiroshi Ishii and Brygg Ullmer. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of CHI '97*, pages 234–241. ACM Press, 1997. 3.2

Gunnar Johansson, Sten Sture Bergstrom, and William Epstein. *Perceiving Event and Objects*. Lawrence Erlbaum, 1994. 4.3, 6.8

Rieko Kadobayashi, Kazushi Nishimoto, and Kenji Mase. Design and evaluation of gesture interface for an immersive virtual walk-through application for exploring cyberspace. In *Third IEEE International Conference on Automatic Face and Gesture Recognition (FG98)*, pages 534–539, 1998. 2.3, 2.4

G. Kaplan. Kinetic disruption of optical texture. *Perception and Psychophysics*, (6):193–198, 1969. 5.7

Antonio Krüger, Andreas Butz, Christian Müller, Christoph Stahl, Rainer Wasinger, Karl-Ernst Steinberg, and Andreas Dirschl. The connected user interface: realizing a personal situated navigation service. In *International Conference on Intelligent User Interfaces*, Proceedings of the 9th international conference on Intelligent user interface, pages 161–168, Funchal, Madeira, Portugal, 2004. 2.1

Esko Kurvinen and Antti Oulasvirta. Towards socially aware pervasive computing: A turntaking approach. In *Second IEEE International Conference on Pervasive Computing and Communications (PerCom'04)*, pages 346–351, Orlando, Florida, 2004. 2.2

Golan Levin and Zachary Lieberman. In-situ speech visualization in real-time interactive installation and performance. In *The 3rd international symposium on Non-photorealistic animation and rendering*, 2004. 3.1.1, 3.9, 3.1.1, 3.1.1, B

Golan Levin and Paul Yarin. Bringing sketching tools to keychain computers with an acceleration-based interface. In *SIGCHI*, Proceedings of ACM SIGCHI 1999. ACM, 1999. 3.1.1, 3.2, 3.11, B

John Maeda. *Design By Numbers*. MIT Press, 1999. 4.4

John Maeda. *Creative Code*. Thames and Hudson, 2004. 4.4

Pattie Maes. Agents that reduce work and information overload. *Communication For the ACM*, Vol.37(No.7), 1994. 3.0.2

Aaron Marcus. Vehicle user interfaces. *Interactions*, (January - February), 2004. 3.0.1

Aaron Marcus and Eugene Chen. Designing the pda of the future. *Interactions*, pages 34–44, 2002. 3.0.2, 3.0.2, 3.3, 5.3, B

D. Scott McCrickard. *Maintaining Information Awareness in a dynamic environment: Assessing Animation as a Communication Mechanism*. PhD thesis, Georgia Institute of Technology, 2000. 4.3

Albert Michotte. *The perception of causality*. Methuen, 1963. 1.3, 4.3, 4.13, 5.3, 5.6, 5.6.2, B

David Molyneaux and Gerd Kortuem. Ubiquitous displays in dynamic environments: Issues and opportunities. In *UBICOMP2004: Workshop on Ubiquitous Display Environments*, Nottingham, UK, 2004. 1.2, 1

N. Hari Narayanan and Mary Hegarty. Communicating dynamic behaviors: Are interactive multimedia presentations better than static mixed-mode presentations? In *Diagrams '00: Proceedings of the First International Conference on Theory and Application of Diagrams*, pages 178–193. Springer-Verlag, 2000. ISBN 3-540-67915-4. 6.8

C. Nass, J. Steuer, and E.R. Tauber. Computers as social actors. In B. Adelsen, S. Dumais, and J. Olson, editors, *Proceedings of ACM Conference on Human Factors in Computing Systems*, pages 72–8. ACM, 1994. 3.0.2

Nicholas Negroponte. *Being Digital*. Hodder and Stoughton, 1995. 1, 7

Les Nelson, Sara Bly, and Thomas Sokoler. Quiet calls: Talking silently on mobile phones. In *SIGCHI'01*, volume 3, pages 174–181, Seattle WA USA, 2001. ACM. 3.5

R.C. Oldfield. *Foreword*. Methuen, 1963. 4.3

J. Pascoe. The stick-e note architecture: extending the interface beyond the user. Proceeding of the 1997 International Conference on Intelligent User Interfaces, pages 261–264, 1997. 2.1, 5.2

H. J. Paulin, M. H. Sørensen, and M. Bødker. Enter the world of ghosts new assisting and entertaining virtual agents. *Paper in progress*, 2004. 1.4

Per Persson, Fredrik Espinoza, Petra Fagerberg, A. Sandin, and R. Coster. Geonotes: A location-based information system for public spaces. In Benyon Hook and Munro, editors, *Readings in Social Navigation of Information Space*. Springer, 2000. 2.3, 2.4

Jean Piaget. *The Childs Conception of the World*. Littlefield Adams Quality Paperbacks, 1929. 1

R. Picard. Affective computing. Technical report, MIT Media Laboratory, November 1995. 3.5

Rosalind Picard and Jennifer Healey. Affective wearables. *Personal Technologies*, 1:231–240, 1997. 2.1

Gopal Pingali, Claudio Pinhanez, Anthony Levas, Rick Kjeldsen, Mark Podlaseck, Han Chen, and Noi Sukaviriya. Steerable interfaces for pervasive computing spaces. In *IEEE International Conference on Pervasive Computing and Communications - PerCom'03*. IEEE Press, 2003. 3.4

Claudio S. Pinhanez. The everywhere displays projector: A device to create ubiquitous graphical interfaces. In *UbiComp '01: Proceedings of the 3rd international conference on Ubiquitous Computing*, pages 315–331. Springer-Verlag, 2001. ISBN 3-540-42614-0. 3.4, 3.4

Edward S. Reed. *Encountering the world : toward an ecological psychology*. Oxford University Press, 1996. 1

Norbert Reithinger, Michael Streit, Valentin Tschernomas, Jan Alexandersson, Tilman Becker, Anselm Blocher, Ralf Engel, Markus Löckelt, Jochen Müller, Norbert Pfleger, and Peter Poller. Smartkom - adaptive and flexible multimodal access to multiple applications. In *ICMI'03 International Conference On Multimodal Interface*, Proceedings of the 5th international conference on Multimodal interfaces, pages 101–108, Vancouver, British Columbia, Canada, 2003. ACM Press. 3.0.2, 5.3

Gerhard Reitmayr and Dieter Schmalstieg. Location based applications for mobile augmented reality. In *Proceedings of the Fourth Australian user interface conference on User interfaces 2003*, volume Volume 18 of *ACM International Conference Proceeding Series*, pages 65 – 73, Adelaide, Australia, 2003. 2.1

George G. Robertson, Stuart K. Card, and Jock D. Mackinlay. Information visualization using 3d interactive animation. *Communication For the ACM*, (vol.36 no. 4), 1993. 4.3

Sverker Runeson. *On Visual Perception of Dynamic Event*. PhD thesis, Uppsala University, 1983. 6.8

Albrecht Schmidt. Embedded information. In *UBICOMP2004: Workshop on Ubiquitous Display Environments*, Nottingham, UK, 2004. 3.3.2

Albrecht Schmidt, Michael Beigl, and Hans-W. Gellersen. There is more to context than location. *Computers and Graphics*, 23(6):893–901, 1999. 2.1

Ben Shneiderman and Pattie Maes. Direct manipulation vs interface agents. *interactions*, 1997. 3.0.2

Gerda J.F. Smets. *Designing for Telepresence: The Delft Virtual Window System*, volume 20, chapter 6, pages 182–207. Lawrence Erlbaum, 1995. 4.3, 4.12, B

Mikkel Holm Sørensen. Enter the world of ghosts. new assisting and entertaining virtual agents. *Available at http://www.itu.dk/people/megel/delcaghosts.doc*, 2004. 5.1

Thad E. Starner. The enigmatic display. *Pervasive Computing*, 2(1):15–18, 2003. 3.0.3, 3.0.3, 3.1, 3.5

John Stasko, Todd Miller, Zachary Pousmann, Christopher Plaue, and Osman Ullah. Personalized peripheral information awareness through information art. In *UBICOMP*, pages 18–35, Nottingham, UK, 2004. Springer. 3.3, 3.3.1

Lucy Suchman. *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge University press, Cambridge, 1987. 2.2

Noi Sukaviriya, Rick Kjeldsen, Claudio Pinhanez, Lijun Tang, Anthony Levas, Gopal Pingali, and Mark Podlaseck. A portable system for anywhere interactions. In *Extended abstracts of the 2004 conference on Human factors and computing systems*, pages 789–790. ACM Press, 2004. ISBN 1-58113-703-6. 3.4

Yasuyuki Sumi, Tameyuki Etani, Sidney Fels, Nicolas Simonet, Kaoru Kobayashi, and Kenji Mase. C-map: Building a context-aware mobile assistant for exhibition tours. In Toru Ishida, editor, *Community Computing and Support Systems: Social Interaction in Networked Communities*, Lecture Notes in Computer Science 1519, pages 137–154. Springer, 1998. 2.3, 2.4

Peter Tarasewich and Christopher Campbell. Privacy and security through pixels. In *UBICOMP2004: Workshop on Ubiquitous Display Environments*, Nottingham, UK, 2004. 3.5

France Telecom. In a bid to develop a new prototype for communicating clothing. france telecom develops an integrated flexible screen to display animated graphics on the wearer. URL http://www.francetelecom.com/en/financials/journalists/press_releases/CP_old/cp040701.html. 3.3.2

Bruce Tognazzini. Principles, techniques, and ethics of stage magic and their application to human interface design. In *INTERCHI'93*, 1993. 3.0.2, 5.2

Andries van Dam. Post-wimp user interfaces. *Commun. ACM*, 40(2):63–67, 1997. ISSN 0001-0782. 2

Roy Want, Andy Hopper, Veronica Falcao, and Jonathan Gibbons. The active badge location system. *ACM Transactions on Information Systems*, Vol. 10 (No.1):91–102, 1992. 2.3

Marc Weiser. The computer for the 21st century. *Scientific American*, 265(3): 94–104, 1991. 1.2

Craig Wisneski, Hiroshi ishii, Andrew Dahley, Matt Gorbett, Scott Brave, Brygg Ullmer, and Paul Yarin. Ambient displays: turning architectural space into an interface between people and digital information. In *First International Workshop on Cooperative Buildings (CoBuild '98)*, Proceedings of the First International Workshop on Cooperative Buildings (CoBuild '98), 1998. 3.3, 3.3

WORDNET. Princeton university., 1997. 1, 2

Ka-Ping Yee. Peephole displays: Pen interaction on spatially aware handheld computers. In *CHI 2003: New Horizons*, volume 5 of *Paper/Demos: Interaction Techniques for Handheld Devices*, Ft. Lauderdale, Florida, USA, 2003. 3.2

List of Figures

1.1 The variety of shapes, forms, and sizes of displays we may encounter in the environment goes from tiny delicate screens to light bulb signs to projections on coarsely textured surfaces. From left to right, large low-resolution LED display on wall, wristwatch display, PDA, and projection on floor. 4
1.2 Overview of physical-world displays. 6
1.3 People may use the temporal dimension of interaction to organize a dynamic environment. In this case, a person uses his car alarm to locate his car in a crowded and noisy parking lot. 8

2.1 Devices in the environment. Features of the environment may influence the range of possible virtual actions. A nearby ticket machine may propose sales without the need for formal introductions. 16
2.2 Bar code equipment. Worker operating engine euipped with bar codes by using scroll buttons on his PDA. 16
2.3 Public virtual notes. Conceptual illustration of notes left by a teacher and available to students outside the classroom. 18
2.4 Triggering by context. System-detected variables trigger a message: if location equals "laboratory" and "temperature" is more than 25 Celsius, then "go to the beach." 19
2.5 The BPN system of changing devices for different situations. From left to right, the desktop computer, the PDA, and the in-car system. 20

2.6 Affective wearable including a ProComp sensing system (upper-left corner) with sensors to measure respiration, galvanic skin response (GVR), and blood volume pressure (BVP), and an electromyogram (EMG) to measure muscle activity. Photo by Jennifer Healey. . . 22
2.7 Ad hoc networks formed by users. Users form spontaneous networks based on the presence of other users. 26
2.8 Can You See Me Now? The abstraction level of representation hides real-world versus virtual-world misalignment and allows "runner" and "player" to share a conceptual understanding. 31

3.1 Interacting by avatar. The self-representation relaying the user's experience. 34
3.2 The Smartakus character. To the right, Smartakus can be seen giving directions via a PDA. Images from www.smartkom.org. . . 35
3.3 The Mob-i personality. The various expressions displayed are, from left: normal, low battery, searching, alert, memory full, "thinking," successful/excited, and reminder. Image (Marcus and Chen, 2002). The design seems vulnerable to loss of detail at changing resolution. 36
3.4 Head-worn displays reveal virtual information about the environment. To the left, normal view; to the right, augmented view. . . 40
3.5 The "virtual-cylinder" interface experience. A user is centered within a cylindrical display of information through which the real world can still be seen. 41
3.6 Using sound as a notifying cue. Sound causes the user to become aware that "College Joe" is online. 43
3.7 Communicating structural location with sound. Sounds may communicate location within a structure. 44
3.8 Using 3D sound to create audible structures. The user may experience structures in space based on a spatial perception of the originating sound. 45
3.9 Hidden Worlds users wearing goggles and "seeing" each others' sounds. Only the shadows on the table are perceivable to outsiders. Image (Levin and Lieberman, 2004). 46

3.10 Messa di Voce installation. The user's voice is visualized by graphics on a large screen. Interacting with the graphics replays the related sound. Image Levin and Lieberman. 47
3.11 Levin and Yarin's. prototype of the acceleration-based interface. The lines move along the x and y axis when the device is tilted. From "Bringing sketching tools to keychain computers with an acceleration-based interface (Levin and Yarin, 1999). 48
3.12 Tangible Bits mapped to digital information, with the objects themselves becoming interfaces. The environment as a whole enables focused attention to the graspable media, as well as awareness in the shape of ambient displays. 49
3.13 Schematic showing the conceptual and perceptual confinement of Window-based interaction with information, as argued by Wisneski and colleagues. The perceptual modality is purely visual, communicating abstract representations of information hierarchies. 51
3.14 A Mondrian-inspired, dynamic "Weather Composition." The position of squares is determined by the relating geographical location of the referring city. Size is determined by temperature, and color by type of weather. Image (Holmquist and Skog, 2003). 55
3.15 Glow Lamp by Elliot and Greenberg. The lamp can rotate its shade on its main axis to change the color of the light emitted. Image (Elliot and Greenberg, 2004). 57
3.16 Ambient Beads. Each bead is suspended by a thin wire from the top of the monitor, and a small motor controls the height of the bead by reeling the line in or out. A continuous motion, or motion at regular-intervals, may be mapped to an approaching event in time. The changed position of a bead could be mapped to a friend's online status: "up" or "down." Image (Elliot and Greenberg, 2004). 58
3.17 Wearable-display image by France Telecom being transferred from a mobile phone. Image Pierre-Emmanuel Rastoin, France Telecom. 59
3.18 Wearable displays by France Telecom. Image Pierre-Emmanuel Rastoin, France Telecom. 60

3.19 The Everywhere Displays projection system works by directing a projection with a mirror. 62
3.20 A carpet display can be used to guide users around. Image (Infineon). 62
3.21 A commercially available lamp like the MAC550 remotely controlled fixture from Martin Light could be used to project information. Image www.martin.com. 63
3.22 A battery of laser pointers could project low-resolution graphics across long distances. 63

4.1 Detail of mosaic from Piazza Amerina, Italy. 70
4.2 Text smeared by anti-aliasing. 72
4.3 Kottke's Silkscreen bitmap font. 72
4.4 Screenshot of Susan Kare's "Woodcut" in MacPaint 1983. Image http://kare.com. 72
4.5 Line-to-pixel conversion. Without anti-aliasing, lines at angles close to the grid lines may result in the "beads" effect. 73
4.6 Closeup of aliasing versus anti-aliasing effect. Top: aliased line. Bottom: anti-aliased line. 73
4.7 Example of dithering techniques. A: no dither, B: diffusion, C: pattern, and D: noise. 74
4.8 Left: BBC logo. Right: The BBC logo converted to the the BBC CEEFAX Teletext system grid. 75
4.9 The two grid systems. Left: conventional. Right: the Nami amorphous grid. Image Heaton. 76
4.10 Three poorly designed symbols shown at different sizes. In the smallest version, the symbols become difficult to tell apart because of the loss of detail. 77
4.11 Game of Life figures, from left: "Traffic lights," "Unix," "Two eaters," and "Small fish." 77
4.12 A single frame from Gerda Smets' "flying bird" at low resolution (Smets, 1995). When all the frames are played as a film, an elegant eagle flies. 78

4.13 One of Michotte's schematics of a visual event. The circle B starts to move with A when they are concentric. Image from (Michotte, 1963). 80
4.14 Magnet TV, 1965. Television and magnet; black-and-white, silent; variable dimensions. Whitney Museum of American Art, New York, Purchase, with funds from Dieter Rosenkranz 86.60a-b. The Estate of Peter Moore/VAGA, NYC. 82
4.15 Picture of the Wooden Mirror as it appeared in *Wired* magazine, Dec. 1999. 83
4.16 Principle of reflecting light from tilted wood chips. 83
4.17 Left: KPN building by Renzo Piano. Right: closeup of display system, with additional standard ticker-style display. 84
4.18 The ITU Interactive Light Walls, by John Maeda, in action. . . . 85

5.1 Conway's "walker" moves from top-left to bottom-right. 91
5.2 Ten Ghost figures are introduced, based on the division of the four species: constant body mass, composite body, physical, and abstract. 92
5.3 "Physical Joe" looks like it's doing exercises. It is supposed to help people remember exercise during the workday. 93
5.4 The "Turner" figure is perceived in terms of object mass. The overall shape doesn't suffer from the heavy diagonal transformation. 93
5.5 The "Struggler" struggles to heave its massive bulk into the air. . 93
5.6 The "Sea" figure reveals that physics are involved in its animated state. The body mass is very flexible, almost fluid. 94
5.7 "Printer Jan" manages print queues and more. 94
5.8 "The Butler" guides people around. He is very busy and consists of 16 frames. 95
5.9 "Thin Lizzy" resides in the cantina and recommends meals. The rotation of elements ensures figure integrity. 95
5.10 The "O-bar" figure. The life cycle of the animation breaks up and joins two bars into a circular figure. The figure survives this transformation well in spite of the fundamental change. 96

5.11 The ”Lovers” figure displays a mutual relationship between two figures that produce the overall impression of a single figure. . . . 96
5.12 The ”L-figure” employs an overlapping technique to produce motion. The resulting animation gives the impression of object constancy by its lack of rigidity of mass. 97
5.13 ”Physical Joe” morphs into a arrow in order to point to the stairs. 97
5.14 Network base station coverage. For the sake of example, the coverage is depicted as perfect circles; in reality coverage varies because of physical barriers. 99
5.15 Calibration grid. Each field within the grid has a measured value used for positioning the user. 99
5.16 The logical layer assigns areas and services to the physical setting. The areas may be constructed independently of physical layout. . 100
5.17 Accessing a Ghost constitutes accessing a specialized service. A Ghost may add liveliness to user conversations by interjecting sensory data or information gathered from the network. In so doing, it may reflect both virtual and physical phenomena. Illustrated here is a Ghost accessing meteorological data from sensors and using them to comment upon the weather. 101
5.18 Information may be displayed on small, inexpensive 8x8-LED displays placed around the environment. 103
5.19 Ghost figures partly cut off and distorted by perspective. The figures are in fact ”L-figure” 5.12 (top left), ”Sea” 5.6 (top right), ”O-bar” 5.10 (bottom left), and ”Lovers” 5.11 (bottom right). . . 104
5.20 The timing of interaction or sound with that of visual cues may produce an experience of causality. 106
5.21 Public-to-public display transitions involve the use of cartoon techniques for discrete communication of the event. The figure in between the displays is invisible until it reaches the new display. . . 108

5.22 A Ghost passing by may be revealed to the user on his PDA while he is engaged in another task, the Ghost Trail causing spatio-temporal displacement of the primary-task interface. To the right, the principle is illustrated by a pattern being broken by the displacement of a circular part of it. To the left, the effect is hardly noticeable as motion is absent in the image. The effect is difficult to replicate with static media such as paper, as the displacement is minimal and is much more readily perceived dynamically thanks to our sensitivity to motion. 109
5.23 Images of an MP3 listing and an area containing a Ghost. 110

6.1 Initials, age, and gender input. 112
6.2 Training on key mappings. 113
6.3 Introduction to the voice-figure tests. 113
6.4 Voice-figure test. In each iteration, one of the figures would display motion that was in synchronization to its voice. This set is in normal perspective. A skewed perspective is shown in 5.19 (see page 104). 114
6.5 The "Who is speaking?" task. 118
6.6 Questionnaire about the figures' appearance. 120
6.7 Subjects were asked to select whether they preferred the blinking or the animated version of the ten figures (the difference cannot be seen in this image). 122

7.1 Screenshot of the Microsoft Windows "My Computer" icon. . . . 126
7.2 Screenshot of some of Susan Kare's classic icons. http://kare.com 127
7.3 The visual design space combining portable displays with wireless positioning and public displays calls for new ways to interact. Bottom-left: PDA, mobile phone, small laptop. Bottom-right: small LED displays, projections, large low-resolution screens. Center: the human user. Top: a virtual service. 131
7.4 Multiple elements challenge the perception of proximal relations. . 132
7.5 Ghost projected onto the floor distorted by the odd angle. 133

7.6 Photo manipulation depicting a Ghost on both watch and distant screen. 134

A.1 "Sea" shown with frame intervals. 137
A.2 The "Butler" shown with frame intervals. 137
A.3 "Lovers" shown with frame intervals. 138
A.4 "Struggler" shown with frame intervals 138
A.5 "O-bar" shown with frame intervals. 138
A.6 "Printer Jan" shown with frame intervals. 138
A.7 "L-figure" shown with frame intervals. 138
A.8 "Turner" shown with frame intervals. 139
A.9 "Physical Joe" shown with frame intervals. 139
A.10 "Thin Lizzy" shown with frame intervals. 139