# Framing Embodiment in General-Purpose Computing

## - a study identifying key components in a multimodel general-purpose computational environment

**By Elisabeth Nesheim**

University of Bergen
Department of Linguistic, Literary and Aesthetic Studies
Digital Culture - MAHF-DIKULT 350
Autumn 2011

# Abstract

The last thirty years have presented us with technology that has had an profound impact on how we produce, socialize with others, and consume culture. Today most of these actions are linked to a computational setup which involves a screen representing our options in two dimensions and a hand-operated controller for manipulating the screen environment, a hardware setup that has not changed considerably the last 50 years. The dominant interface for personal computers—the graphical user interface—is highly ocularcentric, where only parts of the body apparatus (eyes and hands) are addressed in the interface directly. As an increasing amount of information, life experience and human contact is channeled through it, the desktop computer system, becomes increasingly inadequate to fully represent these actions. Any prosthesis added to, or used in conjunction with the body, and any part of the sensory apparatus neglected, will define our interaction with information. Information gathered by the somesthetic—the touch and proprioceptic senses—constitute a significant component in the way we form hypotheses about  what an object is, and how it can be manipulated. By addressing the somesthetic senses in computer interfaces, we can achieve richer and more intuitive interactive experiences.

This paper aims to identify the key components of a general purpose computational environment that foreground multimodal interaction by 1) investigating the significant qualities of the somesthetic senses from a phenomenological and neurophysiological point of view, 2) pointing to successful principles of human computer interaction (coupling), and tools for designing embodied interactions (physical metaphors, interface agents, affordances, and visual and haptic feedback), 3) evaluating the components of current mobile phone technology, surface computing, responsive environments, and wearable computing.

Strategies and plans of dominant technology companies strongly influence what interfaces and devices are available via the commercial market, turning many of us into passive user accepting the default setup made available to us. But if we can move beyond current ideas of what a computer is, re-invent and retell the stories of what we want living with a computer to be like, users are in a  unique position to front and engage discussions that influence artists, programmers, developers and engineers into trying something new.

# Contents

# Introduction

During the last thirty years we have witnessed a dramatic change in how we produce, socialize with others, and consume culture. Today most of these actions are linked to a setup which involves a screen representing our options in two dimensions and a hand-operated controller for manipulating the screen environment. Besides being a work environment, the personal computer is our library, our communication channel, and entertainment console. It is our general problem-solving environment.

The general problem solving environment offered by the desktop stands in great contrast to specific computers or machines where the interface is designed to allow a for a particular activity or to access certain features, as is the case for many game consoles, or many electronic art installations. The Wii is great for moving in a real space to play virtual tennis, but not great for writing a thesis. Any machine's features and functions can be presented in many ways. In early HCI design the functionality of a machine was thought out and implemented before a user interface was designed and set in place. This led to many awkward and to hard-to-use machines that almost demanded a direct access to the developers intentions and ideas, in order to operate them. Increasingly designers have relied on interface metaphors to establish a common ground between machine functionality and possible, often by using real-world metaphors to represent the computer environment. And the general-purpose computer was framed within a very particular: the desktop metaphor, a frame which holds today, forty years after its introduction. The hardware design for the personal computer has thus remained static in its form and setup, despite the fact that more and more information, life experience and human contact is administrated though it.

As of today, many people in the Western world spend more than 6 hours a day in front of this screen, mouse and keyboard setup. And although the activity-level seem high, the body is close to motionless. The two-dimensional representation offered by a screen is also in great contrast with the three dimensional life we live. With so many aspects of human activity linked to one interface that prioritizes the sense of sight, I am compelled to ask how this affects us, especially considering our sense of self and the sense of body.

With these concerns in mind, two questions formed the starting point for my research which has resulted in this thesis. What are the effects of involving more of the human senses (especially the touch and proprioceptive senses) in human-computer interaction (multimodal interaction), and basing interface design on a thorough understanding of human nature and senses? And secondly, is it possible to successfully implement these features in a general-purpose computer design?

The thesis body is divided in five chapters. The first chapter provides a historical introduction to the main computer interfaces that have have led to the dominating interface for personal computing today. It addresses the physical computing environment of the first general purpose computers of the '40s, the development of the mainframe computer and terminal in the '50s and early '60s, before introducing the invention of the mouse and the introduction of the graphical user interface. Finally, it addresses the coupling of the graphical user interface with the desktop metaphor, and the reasons for the persistence of this particular interface.

The second chapter investigates the challenges connected to the graphical user interface and the uncritical use of metaphors in interface design. It particularly points to the limits of ocularscentric interfaces, in terms of accommodating a range of human tasks, as a limited part of our sensory apparatus is actively involved and engaged in the mode of interaction.

The third chapter consist of two main sections, the first discusses the significance of the somesthetic, touch and proprioceptic senses, and the potential benefits of addressing them directly in human-computer interaction. It particularly investigates new haptic and sensor technology employed in mixed reality application. The second section of the addresses the impact of technology, and how interfaces foregrounding only parts of our sensory apparatus affects who we are, what we find important, and finally—what we can experience.
It emphasized the importance of user agency is appropriating new technology.

The fourth chapter gives the reader an insight into dominant design theory, and principles of human machine interaction that has emerged over the last five decades, that has been employed, with varied success, in dominating computer designs. It presents several tools for designing embodied interactions, before arriving at a set of criteria for evaluating four emerging computational devices and environments, alternatives to the desktop computer, that each in their own promote the somesthetic senses.

The final chapter is a case study evaluating four different classes of interface designs that have surfaced during the last decade years, ranging from prototypes, special case interfaces to commercially available products. The aim of the study is to identify key interface components for a general-purpose multimodal computer. In each case the general characteristics and the computational components are evaluated in terms of form factor, mobility, the role of the display, as well as how tactile affordances and haptic feedback are incorporated in the system. The chapter concludes with an outline for what a future personal multimodal computer might look like.

### CHAPTER 1

# An Introductory History of Interfaces

## 1.1 The General Purpose Machine

### The birth of the general-purpose machine

General-purpose computers are—as opposed to specialized computers hardwired to perform specific tasks—reprogrammable, i.e. their base set of computational operations can be configured or programmed to solve a range of tasks.

Charles Babbage designed the first automatic machine, a mechanical calculator named The Difference Engine, in the 1830s. It could be programmed to perform a range of numerical calculations. Parts of the Difference Engine no.1 were build in 1832, but it was never completed. Over the next 10 years, Babbage designed The Difference Engine no. 2 that offered the same computational power with considerably fewer parts[1]. Babbage's final computer design, the Analytical Engine—although never completed in his lifetime, extends the abilities of the Difference Engine. Besides being automatic in its operation, it is probably the first mechanical general-purpose machine, as it used punch cards to program different operations (London Science Museum 2011). The designs of Babbage suggested a new way of thinking about machines and what they could do. The appeal of a machine that could do more than one thing, that could be reconfigured and customized to best centered in on, and solve a task—was far reaching. However, the numerous parts that made up a Babbage machine, and the cost related to making them, as well as the engineering involved in building them, made them rather particular and non-general.

We need to move another 100 years into the future before the sketch of a true general-purpose machine is unveiled. In 1937, Alan Turing proposed a concept of an electronic general-purpose computer, simply called the Turing machine. This machine was never built, only theorized, and within it lay the conceptual framework for the CPU (Central Processing Unit), the operational brain of any computer today. The Turing machine could perform any computational task with simple, operational steps—given a proper algorithmic program, and time. The simplicity and generality of this suggested machine inspired the building of various influential machines the coming decade.

---

[1] The Difference Engine no.2 , minus the printing function, was first build in 1991 by London Science Museum, and is now part of their their fixed exhibition. The machine consisted of 4000 parts and weighted 2,6 tonnes.

## Physical computing in the 1940s

During the 1940s several general-purpose machines were designed and built, with funding primarily coming from military programs of various nations involved in the Second World War.

The ENIAC (Electronic Numerical Integrator and Computer), developed at University of Pennsylvania's Moore School of Electrical Engineering, is considered the first electronic general-purpose computer—and was powered up for the first time in 1947. The ENIAC could perform a range computational tasks, and involved around half a dozen people when operated, see *figures* 1.1 and 1.2.

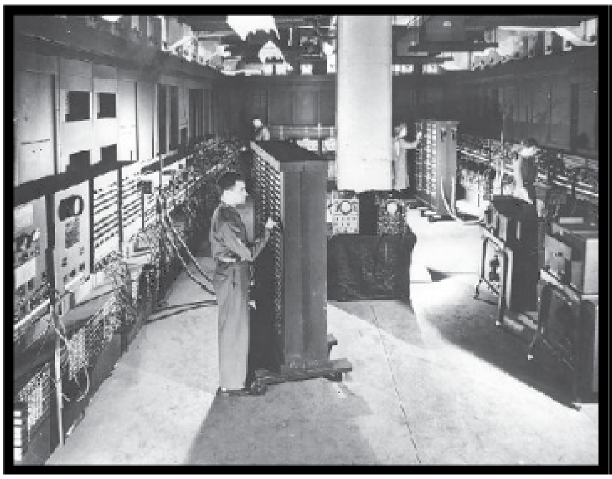

Fig. 1.1. "U.S. Army Photo," courtesy of Harold Breaux. Photograph of the ENIAC while still at the Moore School. Soldier at foreground function table: CPL Irwin Goldstine. Source: Wikimedia Commons

Fig 1.2: ENIAC inventors Eckert and Mauchly (front of the image). Courtesy of the Computer History Museum.

The German Z3, designed by Konrad Zuse and operational in 1941—was the first electromechanical general-purpose computer (fig. 1.3). And a third type of machines—the British Colossus machines—operational in 1943/1944 , were programmable, but disputed as general-purpose machines in the sense that they specially designed for code-breaking (fig 1.4).

A shared interface feature of the giant brains of the 1940s was that the programming and control of the machines was done through physical wiring and rewiring of the machine, by plugging/unplugging cables, flipping levers and pushing buttons. These actions represented the implementing of programs planned and written down in advance. The mere size of the machines demanded that people *moved around* when operating them, and often more than one person was needed at the time to operate it.

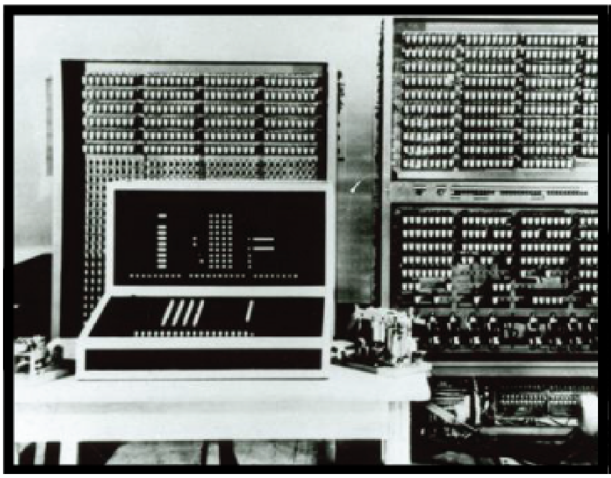

Fig. 1.3 The Zuse Z3 computer. Courtesy of Computer History Museum.

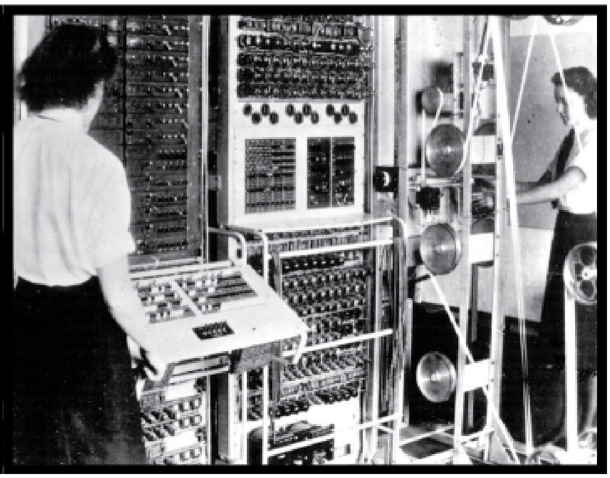

Fig.1.4 The Colossus at Bletchley Park. Courtesy of Computer History Museum.

## Setting up a work station: the computer terminal

In the decade that followed the introduction of general-purpose machines, extensive effort was put in making machines smaller, more reliable and easier to operate. A significant change followed in hardware interface design. From having a group people interacting with the machine from several vantage points in a room, computer terminals that functioned as the operational connector to the overall computer system, e.g. the mainframe, entered the set-up. The first terminals were electromechanical teletypewriters with keyboards and printer display, later replacing the printer with a screen to visualize the input-output activity. The earliest screen-based computer terminals were text based, letting the user interact via a command line interface, before the introduction of terminals that supported a graphical display of information (fig.1.5).

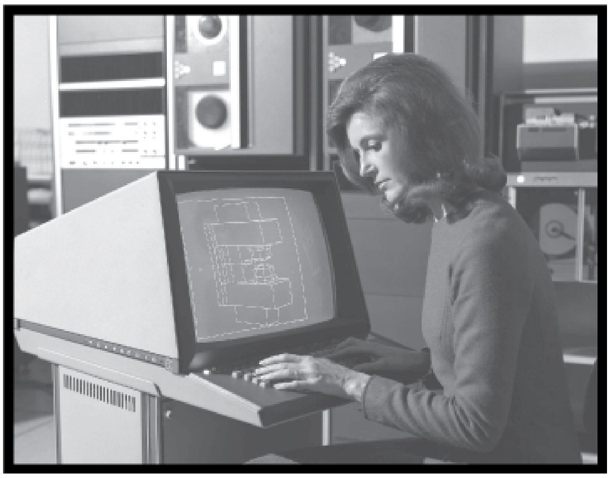

Fig. 1.5 Woman operating the Tektronix 4010, one of the first computer terminals with screens that supported both text and graphics. Courtesy of Rutherford Appleton Laboratory, and the Science and Technology Facilities Council (STFC).

## Graphical displays and miniaturization—the building of a personal computer

The '60s revolutionized how we interact with computers. Similar ideas of what a proper computer system should consist of popped up among several computer scientists, engineers and researchers. With the introduction of graphical displays, windowed screens, different types of pointing devices that allowed for direct or indirect manipulation of elements on the screen and standardization of key functionality on terminal keyboards (QWERTY), the world was presented with the building blocks for a new paradigm of computer systems and user interfaces. The next step was finding the best way to combine them into a consistent whole. The '60s and the '70s gave birth to a range of system designs, many of them never leaving the laboratory of the inventor, while others prevailed. This, along with the ongoing miniaturization of workstations, eventually led to the emergence of different standalone minicomputer systems (such as the IMB 2250, IMLAC-PDS-1 etc.)—and the 2011 edition of the personal computer that I use to write my thesis on.

Of the many possible implementations of a personal computer system, there are two features that, after their introduction, came to dominate the design of personal computer systems for years to come. The invention of the mouse, conceptualized by Douglas Engelbart in 1963, and secondly, the framework for the graphical user interface designed by Alan Kay in the late '60s and early '70s. This framework, which gave birth to the desktop metaphor, is still haunting our screens today.

## Developing the pointing device—Engelbart’s mouse

### NLS—A Collaborative Communication System

In the early sixties Douglas Engelbart and his research team developed the On-Line System (NLS) at the Augmentation Research Center (ARC), part of the Stanford Research Institute. The NLS was the culmination of ARC’s innovations on hypertext, groupware and window interfaces. NLS was a mainframe computer system that could be accessed from several (up to 16) workstations simultaneously making real-time collaboration possible. The design and conceptual framework of the NLS was inspired by American engineer Vannevar Bush’s theoretical Memex system, first outlined in Bush’s influential article “As We May Think” from 1945. The guiding principle of the Memex system was that of interlinked information and functionality—giving Bush the position as the father of hypertext—and Douglas Engelbart food for thought. In his renowned article “Augmenting The Human Intelllect: A conceptual Framework,” Engelbart describes his vision and current implementation of an complex computer system that has the potential of extending man as it increasing “human intellectual effectiveness" by optimizing the ways we gather and process information. He writes: "The entire effect of an individual on the world stems essentially from what he can transmit to the world through his limited motor channels. This in turn is based on information received from the outside world through limited sensory channels; on information, drives, and needs generated within him; and on his processing of that information" (Engelbart 1962). With his On-Line System, Engelbart sought to transform man into an "augmented architect" because, the reach of the computer is not only to solve mathematical problems, or compute number and calculations. "[...] the computer has many other capabilities for manipulating and displaying information that can be of significant benefit to the human in nonmathematical processes of planning, organizing, studying, etc. Every person who does his thinking with symbolized concepts (whether in the form of the English language, pictographs, formal logic, or mathematics) should be able to benefit significantly" (Engelbart, 1962). And any architect must have the proper tools to form, control, and shape his environment.

**Engelbart's mouse**

In 1963 Douglas Engelbart, in collaboration with engineer Bill English, prototyped the first mouse, to be used as a part of NLS. After having run tests with pointing devices controlled by the knee, head and foot, Engelbart landed on a design that involved a hand size box with a vertical and a horizontal wheel inside to track position, in addition to three different buttons to indicate possible actions once in a position, see *figure* 1.6. He describes himself being inspired by an existing pointing device, the light pen—and although not stated specifically, the trackball pointing device invented 11 years prior by Canadian engineers Tom Cranston and Fred Longstaff (Akass, 2001, 24-25), must also have had some influence (fig. 1.7).

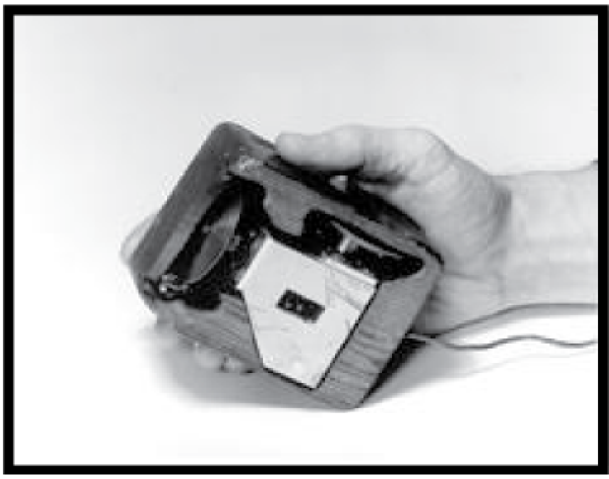

Fig. 1.6 Engelbart holding the first mouse, prototype. Courtesy of SRI International and Stanford Special Collections.

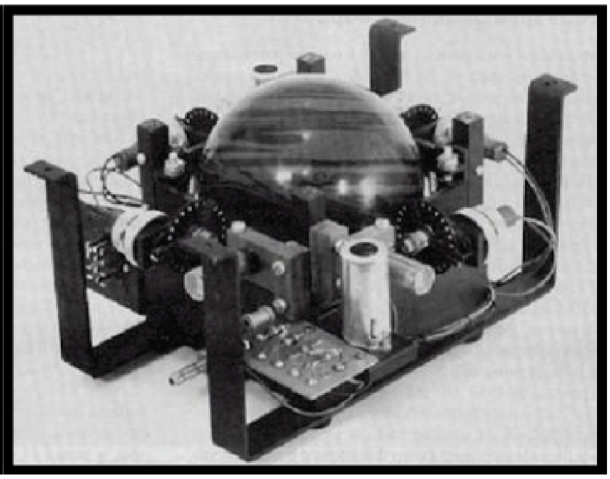

Fig. 1.7 The first trackball device invented by Tom Cranson and Fred Longstaff in 1952 as part of the computer system DATAR for the Canadian Navy.

The mouse was chosen as the preferred pointing device because it showed to be the most user-friendly, the device taking less time to learn—and maybe more importantly the most *precise* of the pointing devices tested at the time. Software architect and multi-touch designer Richard Monson-Haefel made an interesting comment on the additional control devices presented as part of Engelbart's NLS, namely the 5 finger keyboard (fig. 1.8). The user operated the mouse with the right hand, and the left hand was reserved for a specially designed chord key set, an interaction model that allowed for wider range of functionality than the mouse/keyboard combination we are familiar with today. Monson-Haefel describes Engelbart's persistent hunt for efficient ways for people to interact with computers, albeit not necessarily the most easy to use as "The Engelbart Dilemma" (Monson-Haefel, 2008). HCI was, in Engelbart's world, not about usability, but about powerful and efficient ways of interacting with machines, even thought the process of learning the interaction method was a painstaking one. And in the end the chord key-set was simply to difficult for most people to operate.

Fig. 1.8 Engelbart's control desk for the NLS, consisting of the 5 finger chord keyset, a keyboard and the 3 button mouse

## Graphical User Interface and the desktop metaphor

Engelbart's On-Line System came with a preliminary graphical user interface (fig. 1.9), but the windowed version utilizing the desktop metaphor that still is defining of most personal computers today, started with the Xerox Alto computer (fig. 1.10) developed at Xerox PARC in 1973, and computer engineer Alan Kay, was its main architect.

In 1968 Alan Kay began his groundbreaking work of creating a user interface that best combined and utilized the inventions of his time. In accordance with several contemporary thinkers, Kay sought interface solutions that didn't presuppose technically skilled operators, as opposed to Engelbart. Computer literacy should not be about understanding the inner workings of a computer, rather it denotes the ability to "access materials and tools created by others [and] generate material and tools for others" (Kay 1990, 193).

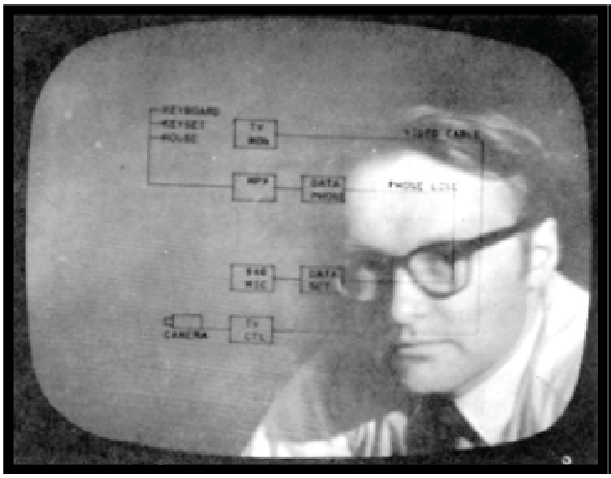

Fig. 1.9 Picture of the NLS screen presented in the article "Augmenting the Human Intellect", with photo text: "Television display obtained by mixing the video signal from remote camera with that from the computer-generated display."

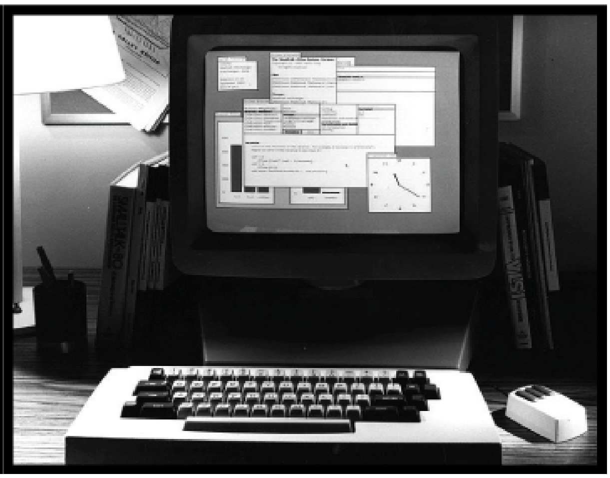

Fig. 1.10 Smalltalk environment in action on the Xerox Alto. Courtesy of Marcin Wichary via the Digibarn Computer Museum.

His proposal was the Dynabook, a notebook equipped with a mouse, keyboard and a graphical display that supported multiple windows. What separated Kay's idea from several other systems of his time was the model for how these elements where keyed together, and how information should be displayed and accessed from the screen. Kay argued that user interface design is tied closely together with learning, and that interaction within a computer environment should be designed as a learning environment.

### Doing with Images makes Symbols

From the research of psychologist Jean Piaget[2] on children's learning phases that includes a kinesthetic, a visual and a symbolic stage, and psychologist Jerome Bruner[3], which claims that a man's cognition is build up by different mentalities, more specifically a native, iconic and symbolic mentality, Kay arrived at his own set of principles for a learning environment and, in extension, a framework for human computer interaction (HCI). Kay condensed his goal for HCI in his slogan: DOING with IMAGES makes SYMBOLS, by proposing a design for a learning environment that includes:

1. a kinesthetic element that allows you to physically manipulate your environment corresponding to the doing mentality,
2. an iconic element, a visual aid involving the possibility to compare and recognize possible actions in accordance to the image mentality, and finally,
3. a symbolic element represented by the learning environment itself that allows for reasoning and learning responding to the symbolic mentality.

---

2. Jean Piaget's works exploring how children learn are presented in *Judgement and Reasoning in the Child* (1928), *The Language and the Thought of the Child* (1926), *The Origins of Intelligence in Children* (1952)

3. Jerome Bruner's *Towards a Theory of Instruction* (1966)

A interface conversation should seek to allow for all the mentalities to co-work, or as Kay puts it, “the best strategy would be to gently force synergy between them in the user interface design.” Alan Kay's first implementation of this model was the setup of the of mouse and keyboard to (doing) and icons (image) in the object-oriented programming environment Smalltalk (abstract reasoning) presented in the computer Xerox Alto (fig. 1.10). And the graphical user interface as we recognize today, was framed. Short later, Merzouga Wilberts, framed these ideas within the concept of WIMP (Window, Icons, Menus and Pointers) that became the default within GUI.

### Modeless interaction and metaphors

A second important feature of Kay’s design was the criteria of no-modes, The main idea is that a user should be able to move from one set of tasks to other within the interaction environment, as well as solving this individual tasks ”without any special termination” (Kay 1990, 197). That is, the user should ideally be able to do what ever he or she desired, from any given starting point. This design decision instilled a need for a metaphor, a way to symbolically describe the environment, so that the user can anticipate which operations it allows. The choice of metaphor was modeled on that of the writing desk, and Kay's graphical user interface was thus populated and structured with the use of virtual objects all recognizable from the physical office world. The screen presented itself as a representation of a desktop that contained endless stacks of papers to be typed on, documents of various kinds and file folders. And although this environment was designed around icons and images taken from the real world, Kay argued that it shouldn't be limited to that. Any metaphor in use should not bring the real-world hassles with it into the virtual environment. The metaphor needs to be magical—and take advantage of the benefits of not having to follow real-world laws of physicality. The graphical user interface, as well as Kay's take on use of metaphors in the Xerox Alto, was ported to fit the more general windowing desktop environment in the Xerox Star computer, and commercialized in the first Macintosh.

The solution for creating fruitful bridges between users and computers became dominated by metaphor design. And after Kay introduced the desktop metaphor as one suggestion, a norm was set. And interesting note that confirms his impact comes from interaction designer and researcher Thomas D. Erickson, while working for Apple in the early 1990s. In his article “Working With Interface Metaphors” from 1990, he presents set of guidelines for interface designers, stating that metaphors should be used in interaction design to ease the interaction between man and machine, and furthermore that metaphors should be based on “real-world events, objects and institutions” (Erickson, 70). Given his position and influence, it is plausible to think that the creation of metaphors based on the above norm, became a default design path within his current working environment, Apple Computer, Inc.

### Commercialization of personal computers

Almost all of the computer systems of the '60s and '70s were built with a specialized operating system that ran with OS-specific software. There was a strong correlation between the imagined operating system (firmware), its intended use, and the design of the hardware. This idea got somewhat lost in the 1980s with the emergence of the commercial personal computer. Now we were presented with one computer for all users and all uses, in a market with several companies competing for the same customers. We see a mainstreaming of hardware user interface design. The Macintosh, followed by an onset of different Windows machines, came in almost identical hardware packages, but with slightly different software functionality. Many of the design ideas of the important computer scientists of the '60s among them Douglas Engelbart, Alan Kay and Ted Nelson, became lost in the commercialization of the personal computer. Fewer and fewer systems were imagined from scratch. Instead new functionality was realized within the GUI and the desktop metaphor paradigm. The commercialization of personal computers throughout the '80s fixed the design convention for years to come in term of how a PC should look and be used.

## 1.2 The Personal Computer

The interfaces of the supercomputers of the '40s surrounded their operators, making it impossible to view the entire machine while working on it. The interaction mode was primarily motor-centric and tactile—through moving between its parts, and plugging, pulling and pressing the various input devices.  Importantly, these machines where equipped with input devices that built upon our cultural and historical use of certain tools. We know from the moment it is identified that a button can pushed, and a lever, pulled. The downside of the machines was the sheer size of the interface—making it hard to get an overview of what actions were performed at any given time, and the current status of the machine, not to speak of the numerous parameters that needed to be set for an operation to take place. This was not the work of an individual, it was a group effort demanding a detailed action plan, prepared in advance and executed in a strict coordinated fashion—discouraging any impulsive exploration of the machine's potential.

### Size does matter

The desire for an overview and a more direct contact with the multitude of actions a user could perform with the general-purpose machine forced the development of workstations connected to a mainframe computer, equipped with informational screens and keyboards, giving the user a single point of access to the functionality and state of the machine. The computer still allowed multi-use, provided through the connected terminals, but the users themselves did not need to plan stringently among each other in order to perform tasks.

This trend followed in the '60s, through a quest for miniaturization, where one key argument, in addition to that of efficiency, seem to be that of ergonomics, reducing the size of a computer to better fit the size of the human operator.  The computer became a tool that a single human being could place (albeit not very often) and operate on at will. The cost was that of mobility—interacting with a computer meant sitting down in front of a screen, and equally important, the computer's input devices—the  keyboard and various pointing devices —demanded of its users a new language of  interaction.

## Escape from TXT

The general move from text-based command line interfaces to graphical ones, opened the door for a whole new generation of computer users. The GUI provided a much-appreciated overview of the content and functionality of the computer. While text based command line interfaces provides a more direct contact with the computer’s underlying firmware—the graphical user interface made navigation through data structures simpler for most users. In fact, the visual representation of data structures gave a sense of overview making many users more daring in their investigation of computer functionality—hence increasing their computer literacy. Arguably, a successful use of metaphors in interface design is of great advantage when creating user-friendly interfaces.

While the command line interface demanded that people recall commands to access the content and functionality of the computer, and not being to accepting of errors—the graphical user interface allowed users to visually recognize icons representing possible user actions and data locations. Another modality was thus employed in helping users navigate and perform their desired actions.

## Framing the computer

With the introduction of the mouse and the graphical user interface, the notion of what a hardware interface should consist of, became fixed. Coupled with the desktop metaphor to guide the development of software, the general-purpose computer turned into the well-known desktop computer. Throughout the '80s this setup reached the commercial marked, branded as a PC—a personal computer—and slowly became a nearly ubiquitous appliance in households in the Western world.

### A personal computer for whom?

The dominating user view of the computer engineers and designers in the early days of the PC is based on the assumption that only was one kind of user. Besides a slight differentiation between what kind of software a particular computer shipped with, users where thought to think the same way when acting when interacting with computers. The early personal computers were fitted to accommodate the needs connected to the profession and work environment of people of the office (accountants, bankers and business men)—the first wave of users who had the financial means to make such an investment.

The following chapter takes a closer a look at the paradigm of the graphical user interface and the challenges connected to an uncritical use of metaphors, particularly the desktop metaphor in interface design, as well as the challenges that have arisen from equating the notion of a versatile computer to that of the desktop computer. And more importantly, it frames the GUI as a particular ocularcentric interface, that along-side virtual reality technologies of the late '90s presupposes that information processing and man-machine interaction is optimized when channeled through the visual sense.

## CHAPTER 2

# The World in My Eyes

## 2.1 GUI and the Prison of the Desktop Metaphor

The desktop metaphor has lured many of its users into a imagining the computer screen as a window to a virtual office desk—with an underlying usability claim that the familiarity with such a surrounding, would ease the transition for any person to become a confident and efficient operator of a computer. Part of this claim is true, as it provided a bridge for the common, non-technical man into the realm of personal computing. On the other hand, it presented a versatile machine in a very particular fashion, framing the conception of what a computer is.

### Metaphors must be magical

Alan Kay saw many challenges with the desktop metaphor as it developed and manifested itself throughout the late '80s and early '90s. His main concern was that a metaphor's prime function is to bridge reality with the virtual in a way that limitations of real-world environments are magically spirited away, by introducing the properties of the virtual. Physical space can be extended or reduced, and information can be collected, grouped, edited, and shared with an ease unprecedented in the physical domain. The intended role of the desktop metaphor, as Kay saw it, was exactly that, to remove physical limitations by introducing the magic of the virtual, because, as he states “one of the most wonderful properties of the computer is that no matter how many dimensions one's information has, a computer representation can always supply at least one more” (Kay 1990, 199).
A metaphor should be one of the building blocks of a user illusion, where one not only immediately understands what real-world object or event a particular icon refers to, but simultaneously how it is magical. And the primary role of the interface designer is to work out the magic.

This has not come to pass. Metaphor generation based on real-life object and situations—rather than magic, became the ruling guideline for the interaction designer in prominent computer production companies throughout the '80s  and as a result, users have had to cope with a representational environment that carries physical limitations within it, to keep the environment realistic or true to its physical counterpart. A classical example is that of the folder structure where a file can only be in one labeled folder at the time (unless it is duplicated), and that connection between folders has to do with proximity in the folder structure. This has been partially mended by the introduction of the short cuts in the Windows

OS or smart folders in Mac OS, but the problem still remains. The physical limitations of the archive have been transported in the virtual realm which initially is free of such limitations. Not only has the average user learned to cope with these limitations, we have, in many ways, accepted them as the default—as the way of the computer.

## Hypermedia as a interface metaphor for user agency

Hypertext pioneer Theodor H. Nelson, a contemporary to Kay, argues that the fixed structure of the GUI, and the desktop environment, is harmful in itself as is hides the true nature of the computer from the user. It presents the computer as a black box—and acts as an abstract layer that, rather then helping a person to work well with computers, removes us from what a computational process really is—and more importantly, what it can be. Just as the properties of real-world objects determine how we can use them, Nelson believes that the desktop metaphor has determined how we think we can use a computer—a limitation imposed by the mindset of class of designers and computer engineers, and not the user.

For Nelson, the goal of interface design is to place the user in control. Keeping the creative force, engagement and our natural ability to dream, should be the utmost concern for the interface designer.  In describing an ideal learning environment for an end user, Nelson proposes a setup that will "[m]otivate the user and let him loose in a wonderful place. Let the student control the sequence, put him in control of interesting and clear material, and make him feel good—comfortable, interested, and autonomous." (Nelson 2003, 313). Nelson argues that our emotional state plays a key role in any mediation. What you see, and what you do has a clear emotional impact on you—it is not only "cognitive structures" that frames our learning.

To understanding the computer is to understand the computer as medium or a multi-medium —to have a "media consciousness." It is not about being a technical expert.  Kay too, envisioned the height of personal computation as consisting of users able create and alter their computational tools without having to become high-end programmers and engineers. But where Kay suggests the extensive use of an iconic language (imagery) to create powerful learning and programming environments, Nelson suggest a hypermedia metaphor, where hypertext and linked media information forms the starting point for gathering, understanding and processing information. As Engelbart, Nelson's proposal is heavily influenced by the conceptual framework of Bush's Memex system.

The ideal hypermedia environment allows the user take any path through it, and form any connection between pieces of information presented within it, based on his desires, dreams and emotions. Nelson argues for strong user agency, which he conceived as improbably within the frames of the desktop. And sure enough, even today, the Window OS form its help section as a FAQ where you only get answers to questions its inventors have predicted you might have. The Apple OS has a more subtle approach, by allowing the user often only one option to perform any specific task.

## Special-purpose metaphor on a general-purpose machine

A third concern comes from leading design theorist Donald Norman who already in the early '90s complained how the desktop metaphor, that its design for particular user actions were not very fitting for a general purpose computer. In his 1991 article “Why Interfaces Don’t Work” he argues that we are where we are because of a “historical accident,“ namely that “we have adapted a general purpose technology to very specialized task while still using general tools” (Norman 1990, 218). Norman looks at interfaces of specialized machines, such as the early video game consoles and household appliances, and attributes their success in the one-to-one relationship between the input device (whether a button, lever, pot or slider) and the action it performs.

Norman claims that ideally “both the interface and the computer would be invisible” and that only “the task that would be visible, the task and the tool being used to accomplish the task” (217). User interfaces should be designed, based on an investigation and understanding of the tasks a user wants to accomplish with a computer. Needless to say, Norman is not too enthusiastic about the one-interface-for-all approach, unless this interface is highly adaptable and modular. This again begs the question whether there can ever be such a thing as a general-purpose computer with a general-purpose user interface, and more to the point, if it is even preferable.

Kay, Nelson and Norman represent the early voices of a growing number of computer researchers and system designers who in the late '80s and early '90s, argued against the single user approach—where all users are the same, and do the same with a computer. From their various viewpoints, they all saw the desktop metaphor as limiting. Direct manipulation of objects, using an iconic language, and framing the mouse, keyboard and screen with a desktop metaphor, was sufficient for managing the significantly few items of information and functions the personal computer shipped with in the early '80s, but became more and more inadequate as the number and types of files grew.

In 1997, Wired writer Steve G. Steinberg looks to computer science professor David Gelernter and his project Lifestreams—a software architecture aiming to replace the desktop metaphor, by presenting electronic documents and email in a streamed time line (Freeman 1997)—hoping he has identified an alternative to desktop vision of personal computing, with which he is not to pleased. “Today, our view of cyberspace is shaped by a 20-year-old metaphor in which files are documents, documents are organized into folders, and all are littered around the flatland known as the desktop” (Steinberg 1997). Lifestream organized all documents, mail and files, by the time they were created and revisioned—the interface metaphor was that of a calendar. The project was in its early stage in 1997. And fairly enough, Steinberg concludes the article with an uncertainty, if it would ever be able to truly compete with the dominating metaphor, particularly because it, as the desktop metaphor was stuck within a very specific hardware setup—a setup shaped by its parents, the engineers, and not the end users. Steinberg's final remark frames one of the reasons why the desktop metaphor became so successful, and the reason why, 15 years after he wrote his article, we are still stuck with it.

## The limits of cognitive engineering

Throughout the '80s and early '90s, cognitive engineering[4] became an established research discipline that sought to profile how human beings operate machines, by looking at how we make decisions and solve problems. The discipline had a significant impact on how user interfaces of personal computers were designed, by assuming that users were rational, and had a concrete plan.

In the *Psychology of Everyday Things*, Donald Norman proposes an approximate model to describe how we engage in actions.

1. We form a goal,
2. we form an intention about how we can reach that goal,
3. we specify an action to take,
4. we execute that action,
5. perceive the state of the world after acting,
6. interpret the state of the world, and finally,
7. we evaluate the outcome.

This is the outline of Norman's model the “Seven Stages of Action”, but he concludes that “[f]or many everyday tasks, goals and intentions are not well specified: they are opportunistic, rather than planned” (Norman 1988, 48). Cognitive engineering does not address the emotional state of the user, nor impulsive acts. And we do perform several tasks with a computer that are equally initiated by by how we feel about something, and sudden spurs of the moment. As computers became networked devices, enabling social interaction between users, much more than rational plans and problem-solving  techniques are involved in our engagement with computers.

---

[4] In 1988 cognitive psychologists David D. Wood and Emily Roth defined cognitive engineering as “an applied cognitive science that draws on the knowledge and techniques of cognitive psychology and related disciplines to provide the foundation for principle-driven design of person-machine systems” (Woods and Roth 1988).

## 2.2 Entering the Cloud

### Interconnected GUI/WIMP computers

The desktop metaphor aside, a major effect of the introduction of the GUI, and a main concern of mine, is that throughout the late '80s and onwards to the early 2000s, computer engineers stopped discussing what kind of hardware interfaces were most suitable for performing various actions and tasks within a computer environment. From the development of the GUI, hardware design of personal computers defaulted as WIMP computers, a hardware setup consisting of a two-dimensional screen utilizing windows, icons and menus, that could be manipulated with a pointing device—the mouse—and a keyboard.
When Tim Berners-Lee announced the protocols for the World Wide Web in 1991, the doors to a whole new world of interconnected users and information were opened to us. The challenge was that, coupled with internet access, in addition to a user's work life, the GUI interface fixed within the desktop now had to accommodates a user's social life: his communication channels, online social activities, and media entertainment.  Even though the vision of the WWW was based on hypertext, interlinked information and websites that users could roam free within, it was implemented within the desktop paradigm. Instead of rethinking what interactive computer system the Internet would be ideal for, it was far too easy to continue down the same path. For Nelson, the implementation of World Wide Web within the desktop, was the ultimate curse:  "I think of the world wide web and XML and cascading style sheets is the ultimate triumph of the typewriter over the author. [...] three fundamental problems today: 1. hierarchical file structures 2. simulation of paper 3. the application prison" (Nelson 2001).

The desktop environment is to this day the prominent setup for which most applications of knowledge production, information retrieval, communication and entertainment are designed. Text editors, graphic programs, managements systems are largely designed to be used by looking at a screen, accessing functionality with a pointing device and keyboard. And despite the desktop going mobile with the introduction of laptops in the early 2000s, allowing us an added flexibility in when and where we can use the computer, most of us still have to perform these actions on a GUI based hardware setup consisting of a screen to visualize our workspace, our social connections, our entertainment media, and information flow, and a mouse or touch pad for pointing and selecting, and a keyboard for entering input. The Internet has challenged the pure desktop metaphor in software design—browsers and media sites allow us to experience information and links between information in ways the desktop metaphor previously discouraged. However, hardware design for the personal computer remains static in its form and setup, despite the fact that more and more information, life experience and human contact is administrated though it. And this is a central concern in my research.

## 2.3 Stepping Into the Screen

The twenty years since the introduction of GUI produced few new hardware interfaces. One noteworthy attempt, albeit unsuccessful in terms of challenging the dominance of the GUI, did come about. In the early '90s an alternative hardware interface for interacting with the virtual and already envisioned by a multitude of science fiction books and movies, saw the light of day: Virtual Reality.

By putting on eye-engulfing displays shaped like glasses, we were spirited into a virtual environment, where we could use our hands to interact with computer functionality or even other users connected to the same simulation. Virtual reality technology was pioneered in the mid-'80s to the early '90s by Scott Fisher through his research at NASA. No desktop, no mouse, no two-dimensional screen was included in the interactive environment of the user. Virtual landscapes were build through extensive programming, and a user could access the environment through the use of huge head-mounted displays (already invented by Ivan Sutherland back in 1968), and gloves for manipulating the virtual objects that inhabited this virtual world.

### The foregrounding of vision

Interacting in VR was interacting in a purely visual environment, and it presumed that visual stimulation was sufficient for allowing the virtual to become real to the user. The actual development of the technology was supported by an extensive selection of science fiction literature where the promise of eye/hand interaction was taken to the extreme. Writers William Gibson described Cyberspace, and Neil Stephenson gave us the Metaverse. Interestingly enough, these authors envisioned their digital domains as interconnected and globally populated virtual worlds, that were as real to us as the physical domain.

Early virtual reality programmers built their visual worlds, based on the research of perception psychologist James J. Gibson and his theory of direct perception presented in *A Theory of Direct Visual Perception* from 1972. Direct perception theory state proposes that visual sense data is transmitted to the brain in full; there is no loss of sense data between it being sensed and processed, and the data flow is unidirectional—from sense object to the sense data processor in the brain.[5]

---

[5] This is in contrast to an *indirect* perception theory, first proposed by psychologist Richard Gregory in his book *The Intelligent Eye* from

Gibson argues that perceptual invariants (texture and linear perception) and affordances (cues in the environment that guide perception) are the cornerstones of perception. And VR programmers sought to translate these kind of aids into the virtual environment to provide direct perception—believing that this would ensure that a virtual landscape could be perceived directly and immediately.

## The limits of virtual reality

We are in the second decade of the 21st century and Cyberspace as previously imagined is still only accessible through fictional sources, for several reasons. Several scare scenarios of how Virtual Reality could be employed to alter and control human beings, were presented in several science fiction feature films and literature throughout the '90s. The idea of direct interaction with the senses (particularly the visual sense) offered the concern of manipulation and control. Who would create the environment, the content presented, and the options available to the user? Others believed that VR would replace common intimacy between people, when most human desires could be accessed without discomfort and stress, in a virtual setting.

These concerns were mostly without merit, as the technology never advanced to a level where it could replace a real-world experience. The first version of virtual reality technologies couldn't keep its promise. The technology used was simply not flexible or precise enough, to create significantly powerful illusions for its users. However, the strongest critics of early VR technology did not considered the latter-day technology as limiting factor for creating life-like virtual experiences. With its one-sided focus on the visual sense and imagining as the prominent gateway to be engulfed a virtual environment, immersion could never take place. Other senses of the body would need to be addressed in order for the experience to be completely mimetic. Still, it is important to realize that early VR technology took one of the first shots at placing the user *within* an environment, in a representation of 3D, instead of *in front* of a 2D representation of interactive environment, and thus in its intention represents a valuable alternative to the desktop systems that dominated and still dominate personal computing. And its successors, the flight simulators and CAVE environments, have proved to be valuable contributions.

---

1970. Gregory argues that perception is 10% sense data and 90% memory. When we perceive, we use past experiences to form hypotheses about what we are perceive. Perception is thus a constructive process.

## 2.4 The Prioritizing of Vision

### Ocularcentrism in VR and GUI

Early VR technologies assumed the eyes to be the prime medium for receiving information from the virtual environment. The virtual reality world of the head mounted displays paid next to no attention to the body of the user. The VR interface assumed that full computer immersion can happen via eyes only, while glove-covered hands do all the talking. Virtual Reality might be most extreme suggestion of eyes/hand computer interaction. But the graphical user interface and setup of the desktop environment are equally based in a learning environment heavily focused on the sense of sight, see *figures* 2.1 and 2.2.

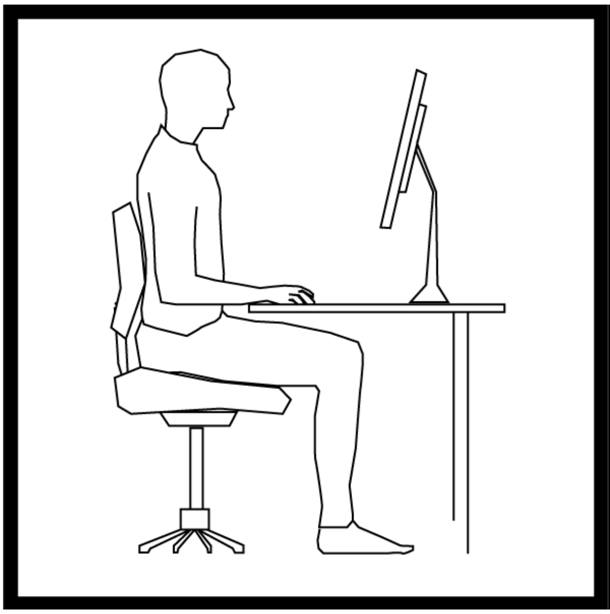

Fig. 2.1: The body posture in the GUI/WIMP hardware setup

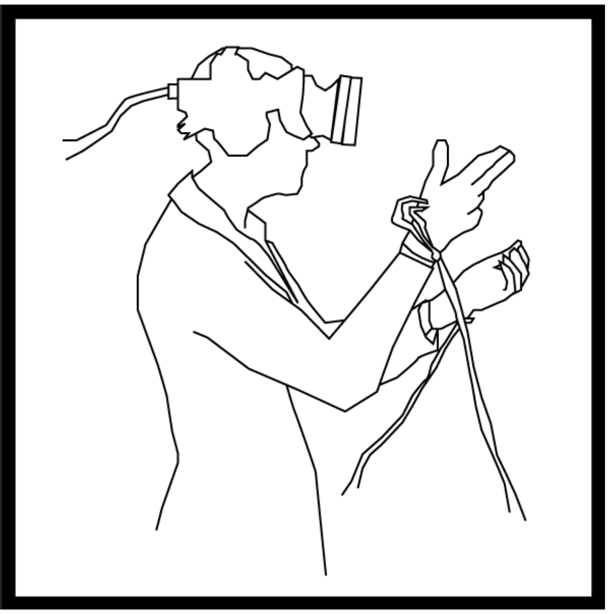

Fig. 2.2: The body posture in early virtual reality environments

Within the GUI/desktop system only particular parts of the body are invoked: Hands are moving, eyes follow moving elements on the screen. When the desktop system is set to accommodate such a wide range of operations it does today, a larger set of experiences are set to be handled by a limited part of our bodies. Furthermore, the metaphor in use suggests that a common ground between computer functionality and real life environment can be communicated through a visually, through a graphical and iconic language.

### Part-time embodiment

Ocularcentric interfaces are not optimized for all of the body. That is not to say that the GUI/VR interfaces claim disembodiment or that they are appealing to the visual sense only. Our entire body is involved in any interaction, but that doesn't mean that we don’t put focus to the parts that playing an explicit role. Our bodies, while interacting with the computer, are receiving a full sensory message, which we do not pay particular attention to.

### Split attention

Visual sense data is insufficient for providing rich interactive experiences, but very capable of directing our mental attention. Just consider the WoW[6] gamer engulfed in hours of engaging play: In-game he is moving between locations, interacting with other players, killing beasts and finding treasures. In the real-world the gamer's body is practically immobile, except for subtle hand and eye movements. It seems that current desktop computing environments are particularly good at engaging our mental attention, but not very good at making us physical attentive. This lead me to question how interface designs and immersive environments foregrounding the visuals sense affect our self-experience, as they certainly lead to disembodied experiences.

## Information is always embodied

Literary critic and theorist Katherine Hayles confronts the assumptions that are held on relationship between mental processing of information and the role of the body. In her book *How We Became Posthuman-Virtual Bodies in Cybernetics, Literature, and Informatics,* she discusses how modern man has transitioned into becoming posthuman, a transition supported by development and implementation of contemporary digital and networked technology. The posthuman era is dominated by a view where information precedes the material, and the body is nothing but an “original prosthesis” that is subject to control and augmentation. And it supports an assumption that mental processes and information can be moved seamlessly between containers. Several science fiction authors have envisioned such information agents, whose personality and intelligence be embedded in different host bodies or hardware, without any distortion or impact. In Greg Egan's book *Schild's Ladder,* bodies are reduced to mere Exoselves, that can be replaced if they are damaged or discarded completely (Egan 2003). That is fiction, but it is also this world-view that has produced the dominant implementation of the computer, where human computer interaction is reduced to a vision-game, that primarily addresses our mental attentiveness.

Hayles counters the assumption that information can move effortlessly and unrestricted between hosts. “Information, like humanity, cannot exist apart from the embodiment that brings it into being as a material entity in the world, and embodiment is always instantiated, local and specific. Embodiment can be destroyed, but it can not be replicated. Once a specific form constituting it is gone, no amount of massaging data will bring it back” (Hayles 1999, 49).

---

[6] The massive multi-player online game World of Warcraft.

A computer's hardware components determine what kind of software can run on it, and ultimately—what kind of information it can process. Human beings are wetware creatures with an operating system, the brain. When neurons fire, and ship sense data from one part of the brain to another, we are witnessing a physical process. Every body instance thus has the opportunity to gather and channel information in its unique way. And any prosthesis added to the body, or any part of the sensory apparatus neglected, will define our interaction with information.

## Looking beyond the sense of sight to the sense of touch and proprioception

Despite the fact that is was not generally successful, the virtual reality paradigm was an inspiration for researchers, developers, designers and new media artists to look for other interface solutions than the GUI.  Tangible computing, wearable technology, ubiquitous computing and augmented reality point to some of the approaches seeking to implement more of the body senses in human computer interaction, commonly labeled mixed-reality technologies. Moving from a paradigm dominated by cognitive engineering, we step into a time for exploring a more embodied metaphor of user interaction.

The following chapter investigates the importance of emphasizing the touch and proprioceptic (motor-sense) senses in human computer interaction, in connection to new digital and sensor technology.

## CHAPTER 3

# Bodies in the Center of Interaction

The body is an extensive sensory apparatus able to gather a vast set of information about its surroundings—through seeing, hearing, tasting, smelling, touching and moving, we interact with the world. We often take our senses for granted. It is only when we lose them (either permanently or by deliberately cloaking them,) that we truly understand how significant each of them are in building our world. It is easy to experience how life-altering the loss of sight is —closing the eyes before navigating a well-known environment, suggests how dependent we are on the visual sense. What is less known is how important the touch sense, and the proprioceptic sense (sense of own position and movement) is in our everyday interaction with the world, partly because it is always active—independent of our attention.

The interfaces in current personal computing technology, as developed over the last thirty years, have paid particular attention to only some of our body senses, while neglecting input from tactile and motor senses. Basing much of our everyday machine interaction on a setup that prioritizes vision over other body senses, affects who we are, what we find important, and finally—what we can experience.

With the onset of readily available digital and sensor technologies, we are presented with new opportunities in experiencing reality, better ways of operating or even entering symbiotic relationships with virtual environments, and lastly, sorting and digesting an increased amount of information. And as we shall see, including touch and proprioception in in our everyday computation is key to exploring these new opportunities.

A second, and perhaps more serious consequence of how technology influence our experience, lies in the opportunity, as well as responsibility in choosing, or even shaping our computational tools to better accommodate the richness of our sensory apparatus.

This chapter consists of two main parts that reflect the above concerns. The first part explores the potential of more body-centric interfaces, and has its point of entry in phenomenological philosophy discussing the role of sight and the critical role of body in mediating new experiences, based on the work of Hans Jonas and Maurice Merleau-Ponty. This is followed by an investigation of the premises of the mixed-reality paradigm, and how our somesthetic sensory system encompassing the touch sense (tactile sensations), and motor sense (proprioceptive sensations), is engaged in technologies that constitute the paradigm. Here we look to haptic and sensor technology currently available on a commercial market, and argue that relevant technology is in place for the creation of successful alternative hardware interfaces that incorporate touch and motor functions more actively. As a historical introduction to the mindset of the mixed reality paradigm, the work of Myron Krueger and his concept of Artificial Reality is presented before addressing Mark B Hansen's view of mixed reality and his proposal for extending the body schema. Finally, the importance of haptic interaction is addressed from a neurophysiological point of view, based on the work of Gabriel Robles-De-La-Torre, as well as contemporary human-computer interaction theory based on the research of Karon MacLean.

The second part of the chapter examines how technology influences our senses, and the importance of user agency is appropriating new technology. This section explores the metaphor of the Cyborg, as a tool for understanding the role of the user, and is based on Donna Haraway's “Cyborg Manifesto,” and the work of performance artist Stelarc.

# 3.1 The Promise of the Body and Body-Centered Interfaces

## The body as significant mediator of new experiences —a phenomenological point of departure

### The role of sight

In his 1954 study “Nobility of Sight”, philosopher Hans Jonas sets out to describe a phenomenology of the senses, a description investigating the assumption of the excellence of sight with roots in Greek philosophy. Jonas assigns three characteristics to the sense of sight to explain its prominent position. Sight is *simultaneous*—when I open my eyes, an image of a juxtaposed now is presented to me, and this image is detached from the objects I look at. In comparison to hearing or touch, sight perception is not sequenced—to see is to partake in “the present as more than the point-experience of a passing now” (Jonas 1954, 513), it is partaking in an extended now. Sight allows us to assign qualities to perceived objects without interacting with them, and regardless of how they change—these assigned qualities linger in our mind's image of them. Secondly, sight *neutralizes* the causality of sense-affection, because we can choose when (by opening and closing the eyelids) and if we want to engage with the seen object. Sight offers a notion of selectivity, which Jonas connects to the ability of being objective—as he writes “from this distinction [between the object in itself and how it affects me] arises the whole idea of *theoria* and theoretical truth.” The perceived image is becoming the object of imagination, a first step in abstraction, where the image can be altered at will, because “[n]o force-experience, no character or impulse or transitive causality enters into the nature of image” (515-516). Sight is therefore the least realistic of the sense, but truly the freest. Finally, sight thrives in the *distance* between the spatial and mental senses, in fact, sight is the only sense that doesn't benefit from proximity to the seen object—according to its capacity for creating an overview of the context, and a point of reference from which to analyze its qualities. Knowing things from a distance is of a great advantage to us (just consider the volume of knowledge acquired with the use of telescopes and microscopes).

Hans Jonas's phenomenological study certainly promotes the qualities of sight in developing our cognitive faculty, which offers as its final remark, that the mind tends to follow the direction of sight. However, Jonas seems to suggest that we need to include both the reflective sense (sight) with the direct experiencing sense (touch), because none of them are sufficient in themselves for grasping reality[7]. No matter the reasoning qualities of our intellect, "[t]ouch is the truest test of reality: It can dispel every suspicion of illusion by grasping the doubtful object and trying its reality in terms of the resistance it offers to my efforts to displace it" (516).

### The first receiver of experiences

The phenomenology of Maurice Merleau-Ponty can be viewed in opposition to the well-established Western paradigm of emphasizing sight and in extension the mind's intellectual interpretation of visual impressions. Merleau-Ponty discuss how an intellectualization of sight is inherent in Western thought, already present in Plato's work, reaching its peak in the philosophy of Descartes, that suggests a final separation of mind and body ("I think, therefore I am").

Merleau-Ponty claims otherwise. He perceives the body as the primary recipient of our sense experiences—the first receiver. Through our long tradition of intellectualizing our sense impressions (and here he uses the Cartesian perspectivalism as a key example), we have lost a direct contact with our senses. In his main work *The Phenomenology of Perception* he states: "by thus remaking contact with the body and with the world we shall rediscover our self, since, perceiving as we do with our body, the body is a natural self and, as it were, the subject of perception" (Merleau-Ponty 2002, 239).

Our bodies are equipped with an extensive set of sensors that allow us to see, smell, hear touch, taste and move in connection with our surroundings. Furthermore we are in a constant dialogue with the objects in the world. Things are more than dead objects that possess certain attributes that can be picked up by our sensory apparatus and deciphered, rather Merleau-Ponty expresses objects to be "complexes" and "unified entities" with which we interact. (Merleau-Ponty 2004, 49). It is the totality of the object we perceive. He continues by stating the objects are never neutral. He states that: "Each one of them [objects] symbolises or recalls a particular way of behaving, provoking in us reactions which are either favourable or

---

[7] Hans Jonas' report on the senses does not mention the proprioceptic sense (sense of movement or position) individually, nor in conjunction with the touch sense.

unfavourable” (48). It is in the meeting between the object and our sense apparatus that we perceive. It is a provoked reaction. Things and their way of appearing are always connected. Perception is a bond between that which sense and that which is being sensed, and this bond is personal.

When perceiving we bring ourselves into the sensing of the object—before our conscious mind have time to formulate and understanding of what we perceive. And this is and important distinction. Having a body with sensors is one thing and perception is the unified impression from this sensor data. And although is seems to that Merleau-Ponty consider this ability a mental one, it is not a conscious one, nor is it located in the intellect/rational human mind.

Through the intellectualization of sight, we have lost contact with the information that our bodies gather about the world. Merleau-Ponty's agenda is clear—he seeks to update our minds through a renewed contact with the body. The way forth in establishing this contact is by becoming more self-aware of the individual conditions that shape our every day perception of the world. And the best source for this is through the interaction with works of art. The reason, Merley-Ponty argues, lies in the artist's ability to present to the world his particular way of perceiving. (Merley-Ponty 2000, 29), and thus giving us a point of reference to familiarize ourselves with our own.

Below, we shall see how contemporary media researcher and theorist, Mark B Hansen, builds on the work of Merleau-Ponty, reasoning that in interacting with particular works based on new digital technologies, the sense data appropriated by the body is revealed to us.

## Extending the body schema through mixed reality

### Emerging technologies and the mixed reality paradigm

From the early 2000s haptic technologies (technologies that utilize a user's sense of touch when interacting with a device), most widely recognized as touch screens embedded in devices ranging from tablets, smart phones and info-boards, to numerous interactive art installations, have become commercially available. In addition, the coupling of computer devices with various sensors (accelerometers, thermometers, light sensors etc), GPS connectivity and triangulation technology, has turned the general-purpose computer into a manifold apparatus able to receive a range of information about its environment.

This development presents a range of opportunities in human computer interaction design, but also a new set of challenges. The GUI and desktop metaphor dominated computer interface design over four decades is, as previously argued, not without flaws. In fact, many of the computer systems designed today utilizing haptic and sensor technology asks users to continue down the same metaphor path. The finger has become the new mouse pointer, and the interaction environment takes place within the familiar two-dimensional screen.

Digital, haptic and sensor technologies presents a core ingredient in the mixed-reality paradigm, a set of user interaction models that seek to extend, augment or enhance reality by introducing elements of the virtual. Ubiquitous systems, augmented reality, wearable computing, tangible computing, haptic media environments, and artificial reality are just some of the names given to what has come to constitute this emerging paradigm of man-machine interaction.

The various mixed-reality systems are differentiated by the degree by which they are visible to their users, whether user actions are tracked and mapped from the *outside* of the body—from the perspective *of* the user's body, or, potentially, from *within* the body—as well as the role haptic or proprioceptic interactions play in the applications. But what all these labels have in common is that qualities of the virtual and digital are introduced into the real environment of the user. The user is present in this, our physical, three-dimensional realm—not a visual representation of a virtual realm, and technology is present to augment, extend, add to, even re-interpret the real environment. An important feature of the applications and interfaces connected to the paradigm, is the position the body proper is given in them, as they utilize and build upon existing knowledge we have of operating the physical world and manipulating physical objects.

Interaction is taking place in real, not simulated environments, where the users—through touch, position, movement and speech—control the virtual elements introduced in the environment. In tangible computing, interaction often involves small physical icons, phicons, that acts as presentation devices of the virtual objects and events, when recognized by the software. Phicons are small, graspable computer devices that recognize virtual markers representing a set of information, and can store and transmit this digital information. Phicons were first demonstrated by Hiroshii Ishii, current head of the Tangible Media Group at MIT Media Lab, and presented in the research paper “Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms” (Ishii and Ullmer 1997).

The mixed reality paradigm has obvious roots in virtual reality technologies of the '90s with the aim of extending the reality of the user with virtual components. But where the VR designers and engineers centered the interface around the visual sense, mixed reality has moved the focus of interaction from the eyes to other body senses, either exclusively or in combination.

One of the earliest examples of the mixed reality paradigm can be found in the work of Myron Krueger and his particular take on interactive three-dimensional environments, and their importance for creating meaningful experiences. Mixed reality is in many ways a re-interpretation of the Krueger's ideas and premises for optimal man-machine interaction.

## Krueger's Artificial Reality—a historical backdrop

The ultimate pioneer of responsive environments is American researcher, media artist and engineer, Myron Krueger (1942—), with his invention of the interactive video projection system VIDEOPLACE in the mid-'70s. In contrast to later virtual reality environments (particularly referring to the HDM accessible worlds of the '90s), Krueger's system operated in real-time without the use of encumbering input devices, and placed the human user in the center of interaction. Real-time movements of the user's body was tracked and projected on a wall where he could interact with various virtual creatures and objects, as well as other human participants.

A set of programs and applications were made for VIDEOPLACE, ranging from simple single player games to multi-user telepresence systems. In 1983, Krueger published his 1974 dissertation which frames his particular take on computer-mediated reality, simply labeled Artifical Reality. The term stems from an argument that AR enhances real-world environments with the introduction of the virtual, but it also plays with a range of contemporary notions of computer-based interactivity of the '80s, namely that such technology was not natural. It would lure its users into a fabricated, fake reality that removed and alienated us from who we really are. Krueger, on the other hand, strongly disagrees with the notion of technology as unnatural or imposed, in fact he sees it as a natural part of human life and culture. “I view technology as the essence of our humanity. An empty hand signals that our anatomy is incomplete until we pick up a tool” (Turner 2002). That doesn't mean that he finds it satisfactory in its current form.

And of importance to our discussion, Krueger argues that man live in a conceptual world consisting of symbols, abstract ideas, myth and language, alongside the physical one—and that an equal cultivation of both worlds, as well as creating strong bridges between them, is vital for human development and meaning creation. But because we have had too strong an emphasis on a set of technologies and tools that primarily cultivate the conceptual world of man, we have created long-lasting disconnections between the body and our minds. “Originally, our conceptual world had no physical or perceptual representation. Later it got worse, reading and writing forced us to immobilize our bodies and to engage only our eyes and brain, rendering the intellect sedentary long before television arrived ” (2002)
Because prominent technologies in place for processing, mediating, representing our

conceptual world have been ocularcentric (vision-centric) in nature, man has diminished contact with his body, as sense data acquired by touch and proprioceptic senses are not emphasized in HCI.

Krueger gives great significance to virtual environments as one of the first promising technologies that can act as vehicle and method making the intellect more agile, as it seeks to incorporate information our bodies have of the world, as well as the mind. His criticism of early VR technology stems from its very continuance of the ocularcentric paradigm, assuming that visual representations of virtuality alone will provide a crossroad between the virtual and real. It neglects to place the whole of the body with its tactility and force in the virtual environment—particularly due to the lack of haptic and proprioceptic feedback, which was one of the key ingredients in his earlier Artificial Reality installations.

The focus on the body in the center of interaction arises from Krueger's fundamental belief that virtual environments and interfaces must build on the real-world experiences of man to provide the necessary starting point for building meaningful paths between the physical and conceptual. It should be a *three-dimensional space*, the *body itself needs to be active* in the interaction—body movement and position is what moves the user experience forward—a feedback loop that continuously makes the user aware of his body-mind, within which we find ourselves everyday. And finally, the interaction must be *real-time* as system lag or poor response times break the connection between the physical and the virtual. The interface itself should ideally be invisible or ubiquitous to the user. The human body is the ultimate interface, and technology is built around its modus operandi.

“In the ultimate interface[...] input should come from our voices and bodies and output should be directed at all our senses. Since we will also interact with each other through computers, the ultimate interface should also be judged by how well it helps us to relate to each other” (Krueger, 1993, 147).

This is the promise of Artificial Reality, but VIDEOSPACE, as a concrete system interpretation of this concept, was build in a defined space, and thus bound the user in terms of when and where to operate in the environment. For Krueger, however, VIDEOSPACE was the best implementation, technologically at the time, for presenting a convincing and reliable virtual environment.

### The qualities of the virtual

Artificial Reality suggests that virtual is not unreal, the virtual is a conceptual framework, a narrative suggestion, into which we can pour both our intellect and sensory information—that can help us form new meaning and structure in our lives. As human beings consist of mind and bodies, both the mental faculty and the body senses should be addressed in virtual environments. Technological applications that utilize the power of the virtual in conjunction with tactile and motor control, seem to be a powerful combination for accessing and applying body knowledge in our everyday lives.

This position is furthered by media theorist Mark B. Hansen, who argues that the body schema can be accessed in new and very significant ways via mixed-reality technologies. Not only can the body, if properly accessed, give us new or complimentary information about our surroundings, the body schema is also democratic, in that it exists before the subject, before interpretation. Hansen, in line with Krueger, believes the reward of including the body when interacting with virtual environments, is significant. They both argue that body should be fully invoked in human computer interaction for the unique reason that the body is able to convey information about our surrounding to us in ways no single sense can. Krueger sees the human body as the ultimate interface, as our bodies seamlessly receive and present a wide set of sense data to the brain, without our conscious knowledge. It is not something we need to control. Interacting with a computer should resemble how we operate our bodies in a real-world environment. For Krueger, that means that particular attention given movement and body gestures, as interaction methods within his responsive environments. Hansen, on his end, addresses the senses of touch and proprioception as primary—and argues that with the rise of new digital technologies, we have the possibility of exploring and unveiling information and knowledge our bodies already possess about the world.

## Exposing and extending the body-schema

### Revealing the body sense

In his book *Bodies in Code*, Hansen states that the body, and not our visual sense nor image of ourselves, is our primary reception and experience organ. More importantly, he argues that mixed reality technologies can mediate and extend our experiences. Part of his argumentation lies in his understanding of the body image versus the body schema, and how a lacking understanding of the body schema has undermined the position of the body in the development of fruitful user interfaces. The phenomenology of Maurice Merleau-Ponty, and his view on the primacy of the body sense, sets the foundations for Hansen's argument.

### Body image versus body schema

The body image is our intellectual imaging of our bodies—it is how I conceive of my body, my thoughts of its potential and limitations, how it relates to and performs when compared to traditional and current cultural body ideals of size, form, smell, flexibility and mobility. It pertains to how I think I look and move, and how I imagine others are perceiving me. It is an intentional interpretation of my embodied self with strong roots in cultural traditions.

However, our bodily sensory apparatus absorbs and processes a vast pool of information about our surroundings without our conscious minds as a guide. I know where my hand is without thinking about it or looking at it. I know instinctively that I am wearing clothes. My interaction with the world promotes an extensive range of haptic and motile experiences, sensations that arises from touch, and from the position and moving of my body. This active information gathering and processing apparatus, is our body schema. Hansen describes it as a "a flexible, plastic, systemic form of distributed agency encompassing what takes place within the boundaries of the body proper (the skin) as well as the entirety of the spatiality of embodied motility" (Hansen 2006, 38).

Hansen sums up the distinction between the body image and body schema with the words of Shaun Gallagher: "In contrast to the intentional (and sometimes conscious) nature of the body image, a body *schema* involves an extraintentional operation carried out prior to or outside of intentional awareness" (Gallagher 1995).

## Extending the body schema

The body schema set us in contact with the *prepersonal—"*the organism-environment coupling operated by our nonconscious, deep embodiment" (Hansen 2006, 20), and by exposing the knowledge and information existing within the prepersonal realm we can enrich, assist and extend conscious reality. Technologies and tools have always been in place to tap into this pool of knowledge, but Hansen believes that new digital technologies play an exceptional role in human evolution, as they represent the development of a new technics. We are now in a particular stage of a technogenesis, our historical co-evolution with technology, in which new digital and virtual technologies can facilitate the creation of media where the body schema is exposed.

For Hansen the promise is great. He claim digital technologies "[c]reate a rich, anonymous "medium" for our enactive co-belonging or "being-with" one another; which thereby [...t]ransforms the agency of collective existence […] from a self-enclosed and primarily cognitive operation to an essentially open, only provisionally bounded, and fundamentally motor, participation" (20). Digital technologies broaden what Hansen calls our sensory commons, which is the domain of sensory information all human beings have access to by virtue of having bodies, it is something all of us share. For me it is unclear, whether he suggests that this sensory commons is a place we actually can communicate with one another, or if it is a shared place all of us have an equal have an access to, but only individually. But, unarguably, Hansen argues that—as previous technology has unlocked and engaged elements of the body-schema in an ongoing technogenesis—mixed reality technologies have the potential to mediate the body schema in ever-new ways.

## Neurophysiological evidence

The importance of activating touch and motor senses in virtual environments has not only been argued from a phenomenological point of view. Several promoting voices are also found in the field of neurophysiology.

### Vital somesthetic feedback

The last five years, a body of neurophysiological research has surfaced, that argues for the importance of somesthetic feedback in virtual, as well as real environments in terms of performance. Gabriel Robles-De-La-Torre, founder of the International Society of Haptics, has made a notable argument. With his background in computer engineering and neuroscience, he refers to studies of patients suffering somesthetic loss (loss of both the touch and proprioceptic/kinesthetic sense). What these research studies show is that sight is a poor compensation for the loss of touch and proprioceptic input, and that people suffering these conditions are severely impaired in terms of acting in real-world environments.
In the case of blindness; auditive, tactile and proprioceptic feedback, can prove to be adequate alternatives for navigating in everyday life—whereas in the case of somesthetic loss—sight has proven to be the most useful, but still a deficient substitute (Robles-De-La-Torre 2006). From the observations, he suggests that the same lack of somesthetic feedback in virtual environments might significantly reduce the performance of its users.
"[A] key lesson is that somesthetic information is critically important for *fast, accurate* [emphasis added] interaction with our environment." So, to fully appreciate an "extended body," which he considers the virtual environments to be, we also need to be able to control it, as we would our physical bodies.

Robles-De-La-Torre argues that the lack of somesthetic feedback in virtual environments can be compared to that of a handicap, and will render the system—the extended body—inadequate for control.

### The potential of multimodal interaction

Research in human computer interaction (HCI) have increasingly considered and evaluated the potential of multimodal interfaces—interfaces that recognizes both input from, and feedback to more than one modality for example a device that recognizes speech commands and touch-gestures as input, and responds with auditory and visual feedback to the user. Multimodal interfaces have long been considered to have great potential in terms of optimizing learning  as they can improve overall perceptual performance, as complementary sense information is combined to identify an object or understanding a task (Helbig and Ernst 2008). Multimodal interfaces have also proved beneficial for the visually impaired (as the visual modality has been key in most general purpose computers).

In recent years multimodal user interfaces have been suggested as a way of coping with the increased amount of data and information we channel via our now-networked computational devices. We have an attention challenge, and a volume challenge, in terms of processing, sorting information available to us—due to an increase in the ubiquity of networked devices. One of the key researchers in this field is Karon MacLean, who is concerned with the design and potential of physical feedback devices to create powerful multisensory-multimodal computer interfaces. As a professor at the Department of Computer Science at the University of British Colombia, he has dedicated his research to "restoring physicality to computer interaction, and to reduce their load on our attention" (MacLean 2011).

He argues that GUI-centered computational environments have, due to the increasing role they play in our day-to-day life and lack of tangible cues, become insufficient devices.

> "...people increasingly do more that one thing at once; because they can and because they now feel they must. Frequently this means their eyes are busy with one task while their ears, hands and/or voice are talking care of something unrelated. For this reason as much as the absence of a large high-resolution screen, having additional information conduits besides vision seems like it might be a useful thing—if our caffeinated brains can handle it" (MacLean 2008).

MacLean sees interfaces incorporating haptic feedback, touch and motor control as increasingly significant and important in terms of deal with our current information load, yet argues we are in the early stages of understanding how to optimize, and set useful standards for the design of multisensory interfaces. Although researchers in the field, agree that there is an obvious benefit from integrating multiple sensory sources in a interaction environment, there is no one ultimate theory of cognition, or complete overview of how our sensory apparatus prioritize the various sense input in any given situation.  Hence, we look at a range of approaches, based on different sets of theories, and most likely a sense of intuition. Maclean concludes: "Touch-derived input plays an unique role in this context [multimodal design], and theories continue to develop on how sensory information is integrated and how conflicting information is resolved. The emerging short answer is that *the task matters"* (2008).

## The key role of embodiment in the formation of selves

Media and performance artist Stelarc, places the body (prominently his own body) in the center of his artistic work. In a recent interview for the Leonardo Electronic Almanac, Stelarc equals consciousness with that of having a body.

> "Consciousness is a characteristic of an operating and interactive body, one that is positioned in a social and cultural history. (To be an intelligent agent, you need to be both embodied and embedded in the world). Insects and animals with a different optical and sensory apparatus would experience the world in diverse ways. Redesigned and re-engineered humans might not only see and move differently, but have an alternate experience of time and space, affecting their interaction with others and the technological terrain they inhabit. We have evolved soft organs to better operate in the biological world. Perhaps now we have to engineer additional organs to better interface and operate with our media and machines" (Aceti 2011, 132)

Stelarc argues, as Hayles before him, that a body, and not a theorizing mind on its own, is a prerequisite for gathering information about the world. He continues that the setup of a body defines what we can experience, and is, in consequence, the capital ingredient in the formation of our identity—our perceived selves. A closer discussion of the work and thoughts of Stelarc will follow below.

We are beings with a body, and the setup of that body sets a frame for what we can experience, and in extension, our consciousness. The content of consciousness in not only shaped by the sensory apparatus, but also our view of this apparatus, and the attention we give it. Sight is a dominating sense, as it is so closely connected to our cognitive faculty and imagination. For centuries Western culture has been infatuated with the sense of sight, as it makes us able to inspect objects and events from the physical world at a will and at a safe distance. For this reason, it has played a crucial role in the establishment of objective reasoning, our ability to form theories about the world we inhabit. Our infatuation with sight is evident in the computer interface design of the last decades—which to a great extent reduces the mode of interaction to that of the eyes and hands. One can argue that we are conditioned by a sight-dominating culture that designed tools in its image. Whether that claim holds, we are in present-day Western societies subject to computer interfaces that primarily engage the cognitive faculty or visual-mental attention of the user.

The role of the body is secondary, and understood as an instrument that can be wielded at will by the controlling mind. But our bodies are constantly participating in our surrounding world, receiving and processing sense data, whether we are consciously aware of it or not.
Information acquired by the touch and proprioceptic senses are vast and unique, and is of crucial importance for our understanding and learning.
The last decade has presented us with a range of technologies that promote a more mobile and tactile mode of interaction, by either foregrounding the somesthetic senses or aligning them in a multimodal setup. Applied within a framework of mixed reality, we can emphasize and extend the reach of our bodies and consequently saturate our pool of experiences.

Directly involving the body schema in our interaction with technology seems purely beneficiary at first, but as with the emergence of any new technology—it is *how* it is implemented for the every-day user that is of the greatest importance. In the following sub-chapter I will investigate how our experience and identity is shaped by our interaction with technology, and the significant role of the informed user in this process. I suggest the metaphor of the Cyborg as a particular helpful tool for understanding what is at stake and what is possible.

## 3.2 The Power of Technology and the Cyborg Liberator

"Science Finds, Industry Applies, Man Adapts"—1933 World's Fair Motto

Any technique or tool underscores some of our body senses to others. Writing puts sensory emphasis on the eyes, and accurate hand control—while beat-mixing with vinyl records emphasizes the tactile and auditory apparatus of the DJ. Throughout history, we find examples of how the different senses of the body are compared and weighted in terms of importance, when confronted with an emerging technology. Technology, tools and techniques are altering our sense ratio.

### Altered sense ratios and the grip of technology

History professor Dr. Robert Jütte, provides an extensive study of the shifts in he prioritization of senses throughout history. His research focuses on the hierarchical position of the different senses over a time span covering Antiquity to the modern day society. Through revisiting the works of Aristotle, Thomas Aquinas, and the philosophy of the Enlightenment to modern day society, he paints a picture of how human agency and our interaction with tools have produced various sense hierarchies throughout time.

> “The hierarchy of senses is both a cultural construction (and therefore based on ideological premises) and a product of the phylogenetic development of the human species (upright physical posture, species-specific increase in the performance of the brain) and the technological changes that have taken place in the course of the process of civilization (displacement of an oral culture by on that is written, the invention of printing, etc.)” (Jütte 2005, 61).

Jütte concludes that the visual sense is dominating our present day society, but does not consider the aspect of power, in terms of a dominating ideology or technology in a given time or society.

With the emergence of media technology a range of researchers have, both from both constructivist and determinist positions, argued for the importance of understanding technology in terms of its impact on identity creation and forming of societies. Technology, in the late media critic Marshall McLuhan's eyes, is a political acteur, with an agenda only revealed to a few. In *Understanding Media* McLuhan argues that each time a new media technology is emerging, most of us are unprepared—we are struck by it, and become temporarily numb. The numbness lowers our natural defense in terms of reflecting upon new experiences. And as new technology settles in—as with photography, electricity, telephony or television—a different sense is foregrounded in our interaction with it. Radio put a particular emphasis on the auditory facilities, whereas television foregrounded the eyes—each media experience equipped with a particular message from its creators and facilitators.

The grip of technology, if not encountered or moderated, is fierce. And the price of being caught in it is high. “Once we have surrendered our senses and nervous system to the private manipulation of those who would try to benefit from taking a lease on our eyes and ears and nerves, we don't really have any rights left. Leasing our eyes and ears and nerves to commercial interests is like handing over the common speech to a private cooperation, or like giving the earth's atmosphere to a company as a monopoly.” (1964, 75). McLuhan is traditionally placed in the determinist camp, but as we shall see, he, along with contemporary researchers more affiliated with a constructivist view, provides us with a way out.

# The Cyborg metaphor
# —connecting identity formation and new emerging technologies

### Haraway's liberating Cyborg

Donna Haraway, contemporary feminist philosopher of science and technology, wrote in 1991 the now-iconic "Cyborg Manifesto." Haraway proposes the myth of the Cyborg to challenge the categories by which we identify ourselves today. The Cyborg has the ability to expose assumptions, values, and beliefs about the current social, technological and political structure —as it is not fixed in any particular position or category.

The Cyborg is both human/animal, man/machine, physical/non-physical—and because of these traits, these qualities it challenges various dualisms (such as self/other, mind/body, culture/nature, male/female, civilized/primitive, reality/appearance, whole/part, agent/resource, maker/made, active/passive, right/wrong, truth/illusion, total/partial, god/man), that have imprinted Western thought, and been the cornerstones in categorizing and dominating women, people of color, animals, nature, and workers as these dualisms enables the "domination of all constituted as others, whose task is to mirror the self" (Haraway 1991, 177). Seeing ourselves as cyborgs we have the possibility to transcend previously conceived borders—and both destroy or give rise to new categories and possible identities.

Haraway believes the Cyborg is a "disassembled and reassembled, postmodern collective and personal self,"—that we must code, constructing a theory and practice that acknowledges this self (163). She points to the rise and importance of communication technologies and bio-technologies as tools that are crucial for identifying our cyborg selves. And it is through technological and scientific discourses about the intention and use of such tools that this identity formation takes place. The Cyborg, a being conscious of its augmentations and in-between-positions, has a unique opportunity to purposefully place itself in the world. It also has the ability to both destroy, as well as giving rise to new categories and possible identities due to its ability to write and report from its standpoint. As the Cyborg let itself explore boundaries, it can give name to these in-between positions, landscapes and realities. The Cyborg can constantly re-tell its story of origin. And most importantly, the Cyborg ability lies within us all.

Haraway, rejecting technological determinism, lets new communication and biotechnologies give birth to a hybrid entity that—no longer blinded by outdated categories—can change existing power structures that now define our social, political and economic reality through conversation, discourse and distribution of opinion. The Cyborg is therefore important for identity creation, and perhaps most importantly, it represents the position of the conscious user willing to reflect on his or her relationship with technology.

### Stelarc's Cyborg—be the Cyborg you want to be

Stelarc is preoccupied with the deficiencies of the body as an interface, and seeks to extend it through modern technologies, previously with robotics and later with genetics. In 2006 he collaborated with a scientist to cultivate a new ear based on his own genetic material, and then committed to surgery to have this third ear attached to his left arm. The ear has functions of a normal ear, and in one stage of the project (although now removed) a microphone was attached at the end of the ear (under the skin of the arm) to transmit the sounds caught by the ear. He states: “a facial feature has been replicated, relocated and will now be rewired for alternate capabilities” (Stelarc 2011). The ear wasn't set in place for purely aesthetic reasons or due to a particular physiological need. In fact, Stelarc reports that it took him years to get someone to perform the surgery, precisely because there was not a clinical need for it. Rather the aim of the project was to extend the evolutionary body in the hopes that “if body was altered it might mean adjusting its awareness” (Stelarc 2011).

Stelarc sees the cyborg metaphor as useful for understanding “what a body is and how a body operates and becomes aware in the world,” and suggests “the Cyborg is the chimera, the recombinant body that performs with mixed realities. Meat, meshed with metal, managing data streams in virtual systems” (Aceti 2011, 136). As Haraway, Stelarc sees the Cyborg as a creature of opportunity, which has the ability to alter his experiences through the use of emerging digital, sensor and bio technologies.

“What is human about the biological body is not only its genetic and physiological repertoire of behavior but that it is an inscribed social and cultural creature that can communicate and collaborate in a multiplicity of media. The body is part of a dynamic and often unstable system of interactivity between other bodies, social institution, cultural conditioning and its instruments and machines. As such the body is not isolated or insulated from modulation and even modification” (130)

Stelarc's view of the body is a challenging one. On the one hand the body apparatus is imperative in the creation of consciousness, suggesting that interfaces neglecting important information feeds from our sensory apparatus are insufficient. Equally, he suggests that the sensory apparatus of every species are different, and thus produce different states of consciousness.  At the same time he doesn't consider the body to be in its evolutionary present, it has failed to evolve in compliance with human imagination and curiosity—and it is through mixed reality technologies, and in particularly bio-technology we can jump the evolutionary ladder of the body. (This rings a bells to Hansen's suggestion that mankind is in an important phase of our technogenesis as discussed above, although not with the same consequences. Where Hansen sees mixed reality technologies as tools for revealing the body schema to us, creating new connection between men, Stelarc seeks to rebuild it.)

Stelarc is an extremist in terms of body modification and Cyborgism, and does not represent the average user in terms of interacting with a powerful, body-enabling interface. But the metaphor of the Cyborg is still useful to understand how we already are conditioned by the technology that surrounds us, and will increasingly be so with the onset of new digital and sensor technology. It is modifying us, extending us, augmenting us. And at the same time—through discourse and conscious choice, the Cyborg represents a position from which the user can challenge the default settings and influencing the coming designs of our everyday computational devices.

### Kittler and the inaccessible logic of media technology

The powerful position of the Cyborg is not recognized by all. Late media philosopher and critique, Friedrich Kittler (1943-2011), is the founder of an objective media theory. The underlying structure of media is not intentions, or feelings—they are purely technical. It is not the human subject that define media reality, but the technical structure of objects. Electric and digital media are able to record information that we are no longer capable of sensing with our body apparatus. We have no real way of measuring or controlling technological media because our senses have been overtaken by them. Kittler no longer sees us as subjects or referents in a technological evolution: media technology has it own underlying structure and logic we are no longer capable of following. He suggest that "[t]he last historical act of writing may well have been the moment when, in the early seventies, Intel engineers laid out some dozen square meters of blueprint paper (64 square meters, in the case of the later 8086) in order to design the hardware architecture of their first integrated microprosessor" (Kittler 1995).

As we no longer have the ability to write, we are no longer in a position of re-writing or altering the media technology that defines our culture. Media interfaces are mere abstractions, control layers that are removing us from the true underlying structure that defines our current culture—hardware.  And even the most technically literate us have a limited ability to alter the inherent logic of media, as they themselves are merely working with an abstraction of a technology able to measure and record data beyond capabilities of a human body.

Kittler's position, although stark and seemingly deterministic in terms of human agency, does help us recognize that the power of technology is far more widespread and ubiquitous that we often give it credit for.

## Becoming a conscious Cyborg

### The artist as a master of sense perception

McLuhan considers only a few of us are able to recognize and evaluate the permanent effects of technology, because it influences our sensory apparatus directly. “The effect of technology do not occur at a level of opinions or concepts, but alter sense ratios or patterns steadily and without any resistance. The serious artist is the only person able to encounter technology with impunity, just because he is an expert aware of the changes in sense perception” (McLuhan 1964, 19). But the few can become many, as the serious artist really is: **“**the man in any field, scientific or humanistic, who grasps the implications of his actions and of new knowledge in his time. He is the man of integral awareness” (72).

Whereas Kittler argues we have a rather fixed technological fate, I find myself in a position of hope, and argue that we need to become conscious and demanding users—serious artists promoting the hardware setups and operating systems that best encompass our everyday computational needs. We can choose what Cyborgs we want to be, and we should choose to be one that better taps into information of the held by the body, instead of one that neglects it.

There is no doubt about the potential of including touch and proprioceptic sense in our personal computing environment, but exactly how it best should be implemented, is as of yet no exact science. The accuracy, responsiveness, and availability of haptic and suitable sensor technologies are increasing, allowing for a surge of experimentation and research in applications. And investigating persistent and successful criteria for tool, usability, and human-computer interaction design, we can stake out a desirable direction for creating a body-centered interface suitable for the individual user.

The following chapter seeks to plow a path through the jungle of design guidelines, principles of human computer interaction that has emerged over the last five decades, to give the reader an insight into the design solutions that have been and still are, for better or worse, dominating the design of our general purpose machines we interact with daily, in hope that we as users can become more conscious in our selection and setup of future personal computational devices.

## CHAPTER 4

# Design Principles for Body-Centered User Interaction

There are volumes of guidelines for designing successful interactions between man and objects. A furniture designer has several considerations to make before his creation hits a market. What is its function? Where should it be used? By whom? What kind of aesthetics are relevant? The same goes for a growing number of interaction designers and engineers presenting new technological tools and devices to a consumer market. What kind of needs do users have? Who are the users? How is its inherent functionality presented to the target group? What is the size of the consumer market? How is it compatible is it with complimentary products?

Human-computer interaction (HCI) is a research and design discipline aimed to assist developers and engineers building interfaces that are intuitive to its users. HCI research has its roots in perception theory, cognitive psychology and communication and tool theory, and is concerned with creating a productive and intuitive entry point to machine functionality. Simply put, HCI seeks to understand how we perceive objects, physical as well as abstract, and how they can be manipulated.

This chapter seeks to give the reader an insight into the choices designers and engineers face when designing successful and intuitive computer interfaces, while outlining dominant design principles in human computer interaction research that have shaped the general-purpose technology we surround ourselves with today. What are the prominent theories and guidelines developed over the last forty years, that not only gave us the graphical user interface but also support the growing number of new alternative interfaces?

## 4.1 Versatility—the General and the Specific

User interfaces vary significantly depending on whether the computer in question is presented as a *general purpose* tool or a *special purpose* tool. The promise of the general-purpose machine is great, at first glance. A one-for-all machine, the Swiss Army Knife of computers, that can help solve any task, anytime, anywhere. It is the ultimate tool—versatile and available for all. And to a great extent this describes the desktop or laptop systems we use today. Whether you are an engineer, a business consultant or a high school student, whether at work, on vacation or at home—this system meets needs of productivity, socializing or entertainment.

The idea of general-purpose machines is powerful, and one currently captured in the Windows, Icon, Mouse, Pointer (WIMP) hardware interface, and a design metaphor, based on an office environment. But as more and more functionality and connectivity is loaded within the same framework, it becomes clear that the original setup is insufficient in representing and accommodating the wide range of use.

Donald Norman, design theorist and usability researcher, points to the difficulty in designing useful general purpose interfaces, due to the lack of powerful general purpose metaphors to bridge the user and the computer. His argument is that user tasks should define the interface, as is the case with special purpose machines. He points to video game consoles of the early '70s and '80s as classical examples of successful machines. They were optimized to perform some very specific actions, and the hardware interface was all about suggesting exactly what actions where available to the user, often by having few and specialized input options, and a direct route to a particular function (starting a computer game was achieved by sticking a game cartridge into the only available slot on the console, and press PLAY/ON.) Most household appliances today, be it a washing machine, dishwasher, heating oven or air conditioning system, are computers. That is, they are CPU-based machines running on a specialized operating system, with a selection of software particular to its domain of activity. They have their starting point as a general-purpose computer, with a specialized hardware interface, which is, with clear exceptions, what makes them so much easier to use than the standardized personal computer. As a consequence, the question that has been bugging designers and engineers over the years is how to insert elements of the general into the specific or visa versa—to create more flexible machines, while still keeping user tasks in mind. So, on one hand, tools should be designed for a specific task to make them easier to use, on the other hand, it is desirable to package functionality to reduce the number of devices in a household or work environment.

This leads us to the hardware/software equation, suggesting that user interfaces drastically change in appearance based on the “ware”-perspective of an computer engineer – that is, whether the user's familiarity with and recognition value of the computer is built via a hardware interface or a software based one. No matter the perspective, both positions hold a particular emphasis on *seamlessness*. A computer should ideally be accessible from any user's starting point, any mode of action—adding an important layer of generality in computer design.

General purpose hardware solutions such as the Windows, Icon, Mouse, Pointer setup for the Graphical User Interface of the '60s were designed with the promise that all software, not matter how specialized, should and could be operated from the same physical control panel. And now in the 2000s we are introduced to portable and scalable software or *general-purpose* *software*, that can be accessed from a range of hardware devices. The fictional Library Computer Access/Retrieval operating system (LCARS) of the Star Trek universe came with a user interface that was primarily based on mode-less software. The user interface is set in software, and scaled to fit any hardware device – from vertical info screens and mobile tablets (the PADD - Personal Access Display Device) to the hallway terminal—where possible tasks and information are presented and accessed in the same way (it should be noted that most Star Trek hardware devices come with a touch screen, as a mayor interaction feature of the user interface.) Today this tendency is probably most noticeable in current application or *app* design—where an API framework dictates the presentation of functionality, and to some extent the mode of interaction. Android Market[8] and Apple's App Store[9] now offer software developed by third-parties, software that has an equal use on a smart phone, a touch tablet or a media player (e.g. iPod Touch). This trend is continued in the design of desktop operating systems, as is the case with Mac App Store for Mac OS X Snow Leopard, or features presented in the coming Windows 8 Metro.

All computers have a hardware aspect and software aspect that needs to be addressed as part of the interaction design solution, and ideally the form factor, hardware user interface, operating system and software applications are equal shareholders in the design process.

## 4.2 Principles of Interaction

### The common ground

In a conversation between two people, several components are put into play. The context is defined by when and where it takes place, the mental and emotional state of the participants as well as their mutual preconceptions about the other. A conversation can turn from engaging to awkward with utterance of a single word that contains a very particular meaning for the one, but for the other means something completely different, or even nothing at all. Sometimes the gap, be it lingual or sociocultural, is just too big—and understanding is replaced by confusion or even frustration. But most often a small explanation is all it takes to bridge the gap, and all is back on track. With every meeting with another person , we update our inner map of what works well at any given time, an ongoing learning process. With small language cues and body expressions we add the necessary information needed to create a common ground for understanding each other.

---

[8] Android Market: https://market.android.com/

[9] App Store: http://www.apple.com/no/mac/app-store/

However, things become slightly more complicated when one of the conversation partners is a machine. The interface, designed and conceptualized by designer and engineers, is the interactive meeting place. Here, functionality of the machine is showcased, and at the same time it must accommodate a range of users wishes depending on their particular context. There is indeed need for a common ground.

Any machine's features and functions can be presented in many ways. In human computer interaction (HCI) design of the late '60s and '70s, machine functionality dictated the user interface, often resulting in hard-to-use machines, where super-users was the ones who best understood the developers intentions and ideas. Later in the '80s, operating system and application software was designed with a preconception that every user did the same with a computer. Human factors and ergonomics research—a still prominent discipline for understanding how man uses tools and how to optimize a work environment for the human anatomy, cognition and orientation—played a key role in perseverance of the WIMP/GUI interaction mode. An extensive volume of research and development have been carried out to provide standards for the form factor and orientation of key parts of the desktop environment we sit in front of today. However, it is a standard of interaction that is based on a generic user, the general man. Even though the view of a user became increasingly diverse throughout the '90s, efforts made to accommodate this diversity were still implemented within the same paradigm.

Brenda Laurel, HCI-researcher, is a strong advocate for utilizing principles from the humanities, especially drama, in man-machine interaction design. Interface design should be centered on *representing possible actions for a user to take*. Both functionality and design should be based in the needs users have, and if not prior to the development of new device or computer environment, the design of the interface be developed simultaneously with back-end functionality. In her book, *Computer as Theater* Laurel defines the common ground as “a shared context for action in which both [man and machine] are agents” where an agent is “one who initiates action” (Laurel 1993, 4).

In agreement with Laurel, Paul Dourish, professor of informatics at the University of California, offers a framework for understanding the constituent of a common ground, that has its starting point in user actions. He extends the concept of a common ground to include the notion of coupling, to explain how meaningful interactions between man and machine take place and how to represent the stage of interaction.

## Embodied actions and meaning

Dourish investigates the relationship between objects in the world, user actions, and meaning. In his book *Where the Action Is*, he shares the notion of embodied interaction as "the creation, manipulation, and sharing of meaning through engaged interaction with artifacts" (Dourish 126). Embodiment is key to any action, and it is intimately connected to meaning. Meaning, according to Dourish, has three main aspects. It can be understood ontologically (as the perceived structure and hierarchical position of individual entities), it has an intersubjective dimension (relating to practice of share of meaning between us) and intentionality (the intended direction of meaning). Dourish believes that is difficult to share meaning by purely ontologically and intersubjective means. In interface design, the intended meaning has to be embedded in the interface by designers. Interface design with a starting point in an ontological approach, is often insufficient, as the interfaces are designed around a designer's or engineer's understanding and structuring of the entities (functionality, features, hardware-components). A user needs to have a similar ontological understanding of these entities to fully understand the system.

Likewise, an interface developer faces many challenges in communicating the constraints and expectations of what a system can for do a user, and equally, predict how a user will come to appropriate the system over time. The exchange of intersubjective meaning in terms of referring to physical objects world is one thing, communicating a belief or expectation is something very different. Dourish finds hope in the third aspect of meaning, intentionality—the directed reference between and object or entity, and another object or entity (its meaning). He argues that computation is always about representation and hence an intentional phenomenon. By introducing a concept of coupling, Dourish seeks to explain how intended meaning can be communicated: "Conceptually, intentionality sets up a relationship between embodied interaction and meaning," and coupling is about making this relationship *effective* (Dourish 138). Coupling is not static bond between an object in the world and a set of actions that can be performed upon it, instead it represents a process of *engagement*, *separation* and *reengagement*. Coupling is an ongoing and active process between a user and an object, and addresses how that relationship is maintained.

Dourish derives his notion of coupling from Heidegger's concept of equipment that describes how tools move from being present-at-hand (the hand and the tool are seen as separate) to become ready-at-hand (the hand the the tool are seen as a single unit, coupled). Both modes are needed for successful and effective use of and interaction with technology. Secondly, we need to be able to freely alternate between modes. And third, it is important whether the tools in question are physical objects or software abstractions. Dourish points to metaphor design, and states that "[o]ne of the best developed uses of coupling in user interfaces does not concern the abstractions in terms of which interactive software is constructed, but rather the abstractions in terms of which the user experience is design(ed) – user interface metaphors" (Dourish 143).

Coupling is a strong concept for understanding the how we can design meaningful relationships between man and objects. It points to the significance of designing objects with a conceptual identity, that equally encompasses the object as a thing in it itself as well as the actions it can assist the user in performing.

## 4.3 Tools for designing embodied actions

As mentioned previously, representation of functionality and possible user actions is key to any computational environment. A user interface sets the stage for interaction between man and machine. It represents an environment, where a user can perform certain tasks and hold cues to how the same user can manipulate this environment according to his/her needs and desires. Metaphor design has dominated user interface design the last decades, but several other design tools have played a significant role creating a common ground where the users and systems are coupled in a meaningful way.

In this section we will take a closer look to metaphor and interface agents as a design tools, as well consider the notion of affordances pursued by design theorist Donald Norman. What is important of all these elements are that they equally relevant whether we apply them in ocularcentric user interfaces or body-centric ones. They both have visual and somesthetic counterparts that can be addressed and employed by the interface designer.

### User interface metaphors

Designers have relied on metaphors to establish a common ground for decades, but employ them very differently. In early interface design, representation was closely connected to a user's manipulation of virtual objects. Metaphor design modeled on the physical desk became a prominent feature in the personal computer, and as discussed in Chapter 2, this move forced users to interact with computers in a very stringent way. Its popularity seems to be founded in an assumption that metaphors founded in real-world objects and events are somewhat easier to understand an access for the most of us, than abstract ones. This notion carries a underlying notion that the computer is a tool, that can be used to represent very specific functionality, rather then conceptualizing the computer as medium that can be formed to fit sets of actions based on a particular user's needs and habits.

Both Dourish and Laurel point to the use of metaphors as a well-proven method to convey an understanding of what a system can do, but these metaphors have to be employed critically, and there is a limit to their use. Laurel is concerned with designing action spaces that conceptualize the overall human-computer interaction, rather than representational interfaces, and is generally critical to the use of metaphors. As she says, it is what “enables you to act within a representation” (Laurel 1993, 21) that is of key importance.
Alan Kay proposes the concept of *user illusion* to better describe the ideal purpose of a metaphor. Metaphors based on real-world objects or event often carry with them limitations that are irrelevant in a virtual environment. The well-applied metaphor helps users recognize how they can perform a certain action within the virtual environment, based on real-world references. But to create a powerful user illusion, the metaphor in question, simultaneously has to make the user aware of what real-world limitations are removed, what actions are made possible in this particular virtual environment. Kay asks: “Should we transfer the paper metaphor so perfectly that the screen is as hard as paper to erase? Clearly not. If it is to be like magical paper, then it is the *magical* part that is all important and that must most strongly attended to in the user interface design” (Kay 1990, 1999). The ability of a metaphor to instill in the user the notion of its virtual qualities, is what Kay calls *magic*, and any successful metaphor should have this magical element.

### Embodied metaphors and gestures

The last decade has given us a range of devices that promotes gestural and movement based interaction.

Myron Krueger is an early advocate of using embodied metaphors in interface design. In all his installations he foregrounds the body by placing the user in the center of a three-dimensional interface, where gestures and movement are what steers the interaction. Krueger is concerned with the *natural information flow* between users and a virtual environment and stresses the importance of an overall physical or embodied metaphor to connect the users body and mind, with possible actions performed within the virtual environment. Krueger is being concrete; the metaphor in question is the body itself. Krueger thus bases his designs on the position of the body, possible movement within the interactive space, and a simple repertoire of hand gestures.

Gestures are becoming more and more common in interface design. Through moving, shaking, pointing, swooping, sliding and pinching we operate our smart phones, access public information terminals and play games. And as we learned how to type touch on QWERTY keyboards, and to navigate the computer screen with a mouse, user and developers will co-develop a new gestural sign language to navigate computer devices and modern virtual environments.

Deciding on a common gestural language is not an easy task. Donald Norman and Jacob Nielsen from the Nielsen Norman Group[10] recently released a status report on gestural interfaces. They are not pleased. First on foremost, they point to the problem of missing established guidelines for use of gestures as an interface control mechanism. Furthermore, they criticize the big computer companies (Apple, Microsoft and Google) and developer communities alike, for ignoring the rich history of human computer interaction research—and instead providing new and inconsistent conventions (Norman and Nielsen 2010), when designing and marketing new gestural interfaces.

## Interface agents

Interface agents are computer programs that help translate and communicate possible option a user can make in various situations. The biggest problem with interface agents are that they don't predict what YOU want, unless they are made to learn from your previous actions in the environment. Simon Penny, artist, curator, and teacher in the field of Digital Art and Technology writes that "[A]ny effective agent interface design project must be concerned with capitalizing on the users' store of metaphors and associations. Agents only work only because they trigger associations in the user" (Penny 1997). This points to the need of the interface agent to be customizable and maybe more importantly, able to adapt to the user's needs and want.  Professor of HCI at the University of Birmingham Russell Beale and researcher Andrew Wood take a closer look to the design of interface agents. They suggest that balanced interfaces promote both users and agent systems as "willing agents", where "[a]gents and their users are more or less equivalent; sometimes one has the leading hand, sometimes the other. Agents view the world from our perspective, and this has to be reflected in the interface, where the agent has to be promoted to a level of equality with the user. This is an interesting move forwards for the desktop metaphor, as it suggests that willing agents, both human and software, can observe, criticize, praise, chide, and learn from each other; a symbiotic relationship between user and software" (Beale & Wood, 1994). Beale and Wood suggest that agent systems need to be equipped with sensors to be able to interact with the same environment the user finds herself in. They need "brains" to form models of the environment they are working within. Additionally, the agent systems need to be able to recognize and interact with other agents, as well as being in an intelligible communication with the user.

---

[10] The Nielsen Norman Group is a consultancy and research company specialized on usability and interaction design. http://www.nngroup.com/

Information from the interface agent can act on different or several modalities based on the design of the system. Through speech, face/eye or gesture recognition, the interface agent receives input from the user, and output can be represented textually and graphically, as well as through auditory or through haptic feedback.

## Affordances—perceiving actions

When we see a button, we instantly recognize that is an object that can be pressed. And the handle on a door, tells us it can be opened (as opposed to recognizing it as a wall with a handle-shaped object fixed on it.) There are several cues, both visual and tactile, that become apparent when perceiving objects, and these cues helps us understand how we can use them. In 1977 psychologist James Gibson coined the term affordance, pointing to the intricate relationship between a perceiving entity and an environment. The affordances of an object or a milieu are the specific properties they offer or provide for a perceiver. Affordances are not objective physical measures, they rather exist relative to the entity in question. Hence the affordances of a door handle to a butterfly will differ from those of a human being. The size of an object determines how being graspable it is, and the surface—whether vertical or horizontal, rigid or soft, steep or mild—decides how well it supports a particular being. Gibson differentiate between the qualities of an object and their affordances, saying that all objects "can be said to have qualities: color, texture, composition, size, shape and features of shape, mass, elasticity, rigidity, and mobility. Orthodox psychology asserts that *we perceive these objects insofar as we discriminate their properties or qualities.* […] But now I suggest that what we perceive when we look at objects are their affordances, not their qualities. […] what the objects affords us is what we normally *pay attention* [added emphasis] to." (Gibson 1986, 134). In addition Gibson argues that affordances of an object does not change even if the need of the perceiver changes. "The observer may or may not perceive or attend to the affordance, according to his needs, but the affordance, being invariant, is always there to be perceived" (139). Hence, Gibson proposes that affordances of objects refer to both their *actual* and *recognized* offerings.

The notion of affordance was later appropriated by design theorist Donald Norman. He threw away the notion of *actual* affordances altogether as they refer to what may be recognized by a user, but not what is recognized. Norman has developed a rich body of design theory utilizing the notion of the *recognized* properties of an object, to understand what actions are perceived by a user as possible when confronted with an object, and how we best can design tools and interfaces that tap into this particular set of knowledge. In his well-know book *The Psychology of Everyday Things,* Norman explains that to best utilize the power of affordances, the designer need to provide a *conceptual model* that allows us to "predict the effects of our actions" (Norman 1988, 13), where the affordances of an object are emphasized or made visible to a user.

Human beings are particular in their way with objects. We not only adapt to how they can be used, we equally modify them, or invent new ones, that optimize our way of living in the world. Or rather, we change the object to what best affords us. The importance of affordance needs to be emphasized in interface design, as it points to what we pay attention to when interacting with an object, how we perceive it as possible to use. It therefore plays a key role in deciding what a user most likely will do when operating an interface. In addition to visual cues, our sensory apparatus opens for adding tactile and motor-centric cues to aid the user in performing his desired tasks. There is a striking difference between operating a virtual keyboard, to a physical one. In both cases, visual cues help us understand how the keyboard can be used, but with the physical keyboard, we have added tactile cues (the slight elevation of each key that tells us how they are separated) that help us type efficiently. The touch sense is directly addressed, which receives sense data about different textures and surfaces, as well as gaps in surfaces, allowing us to distinguish one object from another. As the QWERTY layout is memorized through practice, the tactile cues in physical keyboards become sufficient for operating it—there is no need to look at the keyboard while typing. This feature is not easily transferred to keyboard representations projected on smooth surfaces, or accessible through touch screens.

## The Importance of Response

A key feature in any interface, whether the mode of interaction is based on the visual sense or somesthetic senses, is that of immediacy. Immediacy is closely connected to response of the system, or its feedback mechanism. Whenever a user performs an action, he should be informed about what has been done, and what the action has resulted in. For Myron Krueger the responsiveness of a system is its ultimate feature: immediacy is about real-time response to real-time actions performed by a user, and *lag* is non-existent in any well-functioning interface. Response is the medium: it is what connects the user and the interface.

In the interface design discipline of *direct manipulation*, launched by Ben Shneiderman in 1983, the notion of immediacy is pursued through direct engagement (no-intermediary representations) with the task at hand, and prompt and incremental feedback. Direct manipulation design has primarily been associated with screen-based interfaces such as the WIMP/GUI, where the feedback system is primarily visual. By allowing the user to directly manipulate objects in the virtual environment, e.g by enabling zooming and scaling objects, as well as strategically placing drop shadows to present depth—virtual objects were acknowledged as objects with certain qualities, rather than representations pointing to set of functionalities. The same principle is applied on websites where the updates are done automatically on the server-side, freeing the user from needing to reload a page for his or her action to be updated on the screen. The user should feel that he is directly manipulating an object, and not a representation of an object. Directness is understood in close relation to the notion of distance, particularly the *gulf of execution* and the *gulf of evaluation* (Hutchins, Holland and Norman 1985). Input and control mechanisms should mapped to specific executive actions, and presentation of output (evaluation through feedback) from the computer system, should be immediate and incremental.

But directness comes at a price: by directly mapping achievable to tasks to specific actions, a computer system is in danger of becoming a special-purpose machine.

As Hutchins, Holland and Norman conclude:

"Direct manipulation systems have both virtues and vices. For instance, the immediacy of feedback and the natural translation of intentions to actions make some tasks easy. The matching of levels of thought to the interface language—semantic directness—increases the ease and power of performing some activities at a potential cost of generality and flexibility" (1985).

**Immediacy in haptic systems**

Karon MacLean advocates the use of several modalities in interface design. He generally points to the tendency of implementing haptic design features with the already laden and ocularcentric GUI user interface. (Examples are the use of force-feedback mouses, or GUI-based operating systems that seek to accommodate both classic WIMP interaction and touch screen interaction, as is the case with Windows 7.) In his research paper "Haptic Interaction Design for Everyday Interfaces" MacLean outlines a guide for the novice designer of haptic interfaces. He defines the haptic sense as comprised two modules: "tactile (sensations arising from stimulus to the skin—heat, pressure, vibration, slip, pain) and proprioceptive (which provides our knowledge of body positions, forces and motions via end organs located in muscles, tendons, and joints)" (MacLean 2008). A computer system utilizing the haptic sense will thus vary whether it targets the touch or the proprioceptic sense. Tactile surfaces activating the touch sense need to be precise and responsive, and the area of interactive control (the surface in contact with our skin) is often localized. Interaction design aimed at the touch sense is therefore focused on an understanding of the sensitivity and separation of touch receptors in the skin, to optimize the signals sent from the body to the interface, and from the interface back to the skin. In case with force feedback systems designed to act on the proprioceptic sense, MacLean points to research revealing that we are faster at proprioceptic sensing (receiving force information,) than proprioceptic control (acting on or manipulating objects.) Our overall proprioceptic system operates in close relation to the touch sense, and their dominance is determined on whether we are exploring an environment (mostly proprioceptic) or handling an object (mostly tactile). Haptic hardware system thus needs to recognize the qualities of our somesthetic senses, to strengthen the signal flow between the user and the system.

## Conceptualizing human-computer interaction

Powerful interfaces, interfaces that define a common ground between man and machine, need to be founded in an overall conceptual model that considers man-machine interaction in its totality. The input and response mechanisms of hardware components as well as software components (operating system and user applications), need to be designed to emphasize the tasks a user want to perform to reach his goals by interacting with the system.

Metaphors have been a popular way to build a conceptual model for how the system is used, but they are often limited in use, due to being inconsistent or only applicable to understanding parts of the overall computer system (a set of applications presented within a specific operating system, accessible through a specific hardware interface). Physical gestures seem particularly promising as they are based in how we operate our bodies and thus can leverage on user intuition. However, viewing the recent development of gestural interfaces, we are in danger of establishing unhealthy or competing conventions of use of gestures that confuse rather than aid the user in reaching his goals. A conceptual model should emphasize the power of affordances. The designer thus need to recognize the visual and tactile affordances an physical or virtual object when building an interactive environment, to best direct the attention of the user towards possible action he can perform within it. Finally, any interactive system greatly benefits from active and precise response to input from the user.

### Standards for haptic interaction

In 2009 ISO (International Organization for Standardization) released “ISO 9241-920:2009—Guidance on tactile and haptic interactions”, and two years later it was followed up by “ISO 9241-910:2011—Framework for tactile and haptic interaction”. The content of the standard is outlined in the research paper “Setting the Standards for Haptic and Tactile Interactions: ISO’s Work” published by the authors and members of the ISO work group. According to the research paper the standard includes a definition of what a tactile/haptic interaction is, when to use it (types of tasks, techniques, strategies), cases to illustrate the varied use, strategies for the mechanical coupling in terms of force/feedback systems, and the overall the effect of haptic systems ( Jan B. F. van Erp, Ki-Uk Kyung, Sebastian Kassner, Jim Carter, Stephen Brewster, Gerhard Weber and Ian Andrew 2010). The paper does not mention what haptic technology or design solutions, that have been evaluated or considered in the setting the framework, for that one need to purchase access to the standard itself.

Setting a standard for creating a common framework and convention for haptic interaction, is a good thing. If the standard is adopted by manufacturers, users will not have to adapt to a range of systems, and it opens the door for mutual compatible design solutions developed by various companies and engineers. Standards are influenced by technology and design features promoted by the current market leaders, but as Donald Norman and Jacob Nielsen state in above-mentioned report, current gestural interfaces developed by marked leaders often ignore established principles of successful user interaction, leaving user to deal with inconsistent solutions (Norman and Nielsen 2010). Equally, a range of inspiring haptic interfaces are presented in electronic art and DIY technology projects. Users, artists, and independent developers have important insight in conceptualizing and operating haptic interfaces that should not be ignored. And my hope is that we are invited to take an active role in any discussion that involves setting fixed criteria for how new emerging technology is developed.

**General-purpose technology and general-purpose interfaces**
This thesis is concerned with the dominating framing (GUI/WIMP) of general-purpose technology. The concern does not only extend to problem of vision-centrism, but additionally the concept of general-purpose interfaces. Can such interfaces ever be designed? This chapter has been focusing on design principles and tools for creating intuitive interfaces, but there seems to be give-and-take relationship between optimizing an interface for specific tasks and actions, versus general use.

The computer is a general-purpose machine that can be shaped to accommodate a number of tasks. Instead of pursuing the quest for a general-purpose interface, I propose the notion of *versatility* to describe the extent of use in personal computing systems.

## 4.4 Evaluating Body-Centric Interfaces Based In General-Purpose Technology

The goal of my research is to acknowledge the potential of utilizing more of the body senses, particularly touch and proprioception, in human-computer interaction, and secondly, to investigate how they can be incorporated in a general-purpose computer system. The mixed reality paradigm has proposed several computational devices and environments that each in their own promote the somesthetic senses. What is important when considering new interfaces that foreground proprioceptive and touch senses to that of sight? Below I suggest four criteria to evaluate body-centric computer designs, and if possible, identi the overall conceptual model employed.

### Versatility

Versatility points to the range of user actions a computer system accommodates. What kinds of tasks does it allow us to perform? How customizable is it? And how does it compare to the many tasks we can perform within the standardized desktop workstation?

### Direction

Each interface directs our attention in their unique way.
I am DJ, and most times my hardware setup consist of two record players, a mixer and a set of vinyls, and I am coupled with this interface through tactile and auditory stimuli. With the introduction of vinyl emulation software (e.g. Traktor Scratch,) I can play back digital tracks via vinyls. With this feature a second interface, a screen-based one, is introduced into the mix. The added bonus of emulating vinyl is both economical and practical (in terms of carrying weight and storage.) However, the price to pay in terms of having to move the center of my attention from the auditory and tactile domain, to start interpreting and responding to the visual information presented on the screen, is high, and can degrade my overall performance. In most screen-based interfaces our attention is directed at the visual representation and interpretation of information. Sight is powerful modality, and interfaces foregrounding sight are generally aimed at directing our attention towards the cognitive facilities. When analyzing body-centric interfaces, the employment of a screen, if any, is therefore of interest in terms of evaluating alternative configurations of the mental and physical attention.

### Mobility

I use the term mobility to point to the portability and form factor of a computer device, and hence when and where we can interact with it. Mobile computing relates to a user's ability to move around, or between locations when operating the device in question. The desktop fixes the user to a desk. Heavy duty laptops are neither considered particularly mobile, rather migratory, whereas smart phones or smaller tablet PCs allow us to interact with the device in any particular location—some even while moving.

### Tactility and proprioception

I seek to identify how tactile and proprioceptive (somesthetic) senses are engaged by pointing to the extent that physical metaphors, tactile affordances, tactile sensations and force feedback mechanism are employed in the interface. How does the interface utilize the tactile knowledge of the user? Are there any embodied metaphors in play? How does the system respond to user? To what extent is moving or the position of the user driving the interaction?

## Towards a conceptual model of body-centric interaction

In the later years an new interface design discipline has emerged that aim to create interfaces that appear natural to the user. Interacting with a computer should be similar to how we interact within a real-world environment. These interfaces, labeled Natural User Interfaces or NUIs, represent a vision of creating interactive settings that accommodate more of the body senses in human-computer interaction.

### Microsoft's NUI definition

While the constructs defining a GUI interface are considered the WIMP elements: Windows, Icons, Menus and Pointers, the NUI interface are build up by different constructs. August de los Reyes, creative director of the Window Platform Core team, argued at a 2008 conference, that the following five constructs or elements needed to be part of any natural user interface: First the environment needs to be *contextual*, the interface should be designed to accommodate the tasks a user want to perform at any given time, secondly, the interface should accommodate *scaffolding*, where the user learn what he can do within the environment, as he is interacting with it. Thirdly, it should accommodate the *direct manipulation* of virtual objects, and not representations of virtual objects. Forth, the performance esthetics should leave users ecstatic about their *experience*. And finally, the interaction should allow user to *extend their real-world* experiences (de los Reyes 2008).

While the two last constructs say more about how Microsoft designers want people to feel about their interactive experiences, the three first constructs points to established conventions in HCI theory. Contextual environments seek to address the user from his point of departure, letting his current needs frame the conversation. Equally, the learning-as-you-go approach, rather than forcing a user to study how the interface works before operating it, is one of the key success factors of specialized machines. Thirdly, direct manipulation removes a representational layer between machine functionality and actual user actions.
Joshua Blake, technical director of Infostrat, a company producing natural user interfaces for the public sector suggests that “NUI is an interface that is designed to reuse existing skills for interacting directly with content” (Rick Barraza et al. 2011). Blake addresses an equally crucial point, namely that interacting with computer should resemble how we operate our bodies in a real-world environment. Current notions of NUI design thus seem to be a suitable starting point for arriving at conceptual model of body-centric interaction, especially if the design features are maintained in both software and hardware aspects of the overall interface. It is the combination of software and hardware that constitutes the rich natural interface.

In the following chapter I will take a close look at four current and upcoming interface designs that can act as body-attentive alternatives to the desktop computer environment of today. I will evaluate them based on their versatility in use, the weight a display is given in user interaction, how mobile they make us, and finally how the somesthetic senses are engaged. From this I will look at the suggested mode of interaction, and see how well it is maintained in the design of the hardware interface, the operating system and the applications presented within it.

CHAPTER 5

# Case study: Identifying Key Interface Components for a General-Purpose Multimodal Computer

In this chapter I look at four different classes of interface designs that have surfaced during the last decade years, ranging from prototypes, special case interfaces to commercially available products. The different interfaces are each optimized for particular tasks and modes of interaction. A common feature shared by all is that they accommodate multimodal interaction. I address the mobile, tactile and proprioceptic qualities of these new and emerging computational environments, and investigate how they can augment or even replace the current personal computing environment based on the Graphical User Interface and WIMP setup. A particular emphasis is given haptic interfaces that downplay the role of the screen as a key component for representing the interaction environment, as such interfaces reduces the cognitive overhead of visually interpreting the functionality within the interface, while focusing on the task at hand.

I will look at the general characteristics, the computational components, of each environment, in terms of form factor, mobility, the role of the display, as well as how tactile affordances and haptic feedback are incorporated in the system. Then I will point to relevant specifics of each of the environments, identifying key interface components for a body-centric general-purpose computer.

## Cases

The user cases are examples of different combinations of mixed-reality technology, each of them interfaces that emphasize the haptic interaction, and optimized for certain user actions—a set of computing needs.

1. **Mobile touch screen devices**
   Smart phones, Media players, Tablet PCs
2. **Responsive surfaces**
   Interactive White Boards, Tangible Bits, Microsoft Surface, Reactable
3. **Responsive environments**
   CAVE, Artificial Reality, Kinect
4. **Wearable computers**
   Sixth Sense

## 5.1 Mobile Touch Screen Devices

Mobile phone technology became the prominent communication technology in the mid '90s. Telephone communication became mobile, and was no longer about connecting users in two different fixed geographical locations, but rather between any geographical location at any given time. Coupled with text messaging functionality, mobile phones replaced landlines—and instantiated an expectation of a user always being available, since he was never more than a phone call away. The devices themselves were hand-held and came with small displays to accommodate text messaging and visualizing incoming and outgoing calls.

One decade later, and functionality packed within mobile phones now included cameras and media players, and to some extent emailing and internet browsing—although the latter two were specialized features usually made available through PDAs (Personal Digital Assistants.) PDAs were successful devices, particularly in the business life, as it, at the time, combined the popular features of two computing environments in one device: it was highly portable, compared to a personal computer, and secondly, it had a better user interface than the average mobile phone for handling email, calendars and scheduling programs, as well as internet browsing, due to the possibility of inputting commands and navigating the system using a stylus via a touch screen interface.

Fast forward another couple of years, and PDAs gradually became superfluous as their features migrated into mobile phone devices. Mobile phones turned into smart phones or multimedia phones, where each device had their unique way of balancing the quality of various features in the device. Some devices came shipped with first-class cameras, whereas others were optimized to sync with your existing email and calendar system, and others, again, were just excellent phones, compatible with many different cellular networks. Telecommunication, instead of being the dominating feature, was now one of many features in a range of microcomputers running on specialized operating systems. Generally, the devices themselves became thinner and smaller, while their displays were increasing in size. The bigger displays accommodated the need of representing the increasingly complex menu and settings systems, to allow a user easier access to the particular features. And gradually, as the displays have grown in size they have been replaced by touch screens, either in combination with keyboards (or number pads), or just by a few control buttons.

### Mobile multi-touch devices

In 2007 Apple introduced the iPhone, the first multi-touch smart phone, as opposed to the earlier single touch displays, that only allowed point-and-click control, or single-touch gestures (e.g. a swipe.) The hardware interface of the iPhone was, with the exception of a few buttons, a single touch screen surface. The second upgrade of the iPhone operating system (mid 2008), included the App Store, and for the first time third party developers were given API access to a mobile phone operating system, opening for a range of new applications instantly compatible with, and easily available to users of the iPhone. The portable media player iPod Touch, launched in 2007, ran on the same OS and came shipped with the close to an identical hardware interface as the iPhone. The iPod Touch devices gained access the App Store just a couple of months later than the iPhone, and from then on the only significant thing separating the two devices are the GPS and phone (GSM modem/SIM) applications.[11] As the smart phone devices and media players have become increasingly versatile in use, a third product, the touch tablet PC, gained popularity. Tablets, bigger in size, can host a more powerful operating system, and a more prominent display—accommodating use that is generally connected to desktop operating systems. A full-size virtual keyboard allows for more efficient typing, and the bigger display optimize reading, and entertainment wise, playing video games and watching movies.

During last two years the consumer market has been booming with portable touch screen computers— either in the form of smart phones, multi-media players or tablet PCs—that share several characteristics:

1. They are highly portable due to weight and size, and can be operated close to anywhere depending on which subset of their features are accessed.

2. The display and the area of operational control are presented on the same surface.
This assembly is of an advantage in the smaller devices, as it is possible to hold and operate the interface with one hand, and comfortably by using both hands. As the devices increase in size, they become more difficult to hold and operate simultaneously, as is apparent with tablet PC's in particular. Since display and interface controls are on the same surface (and not split between a horizontal and vertical surface), extensive use of the device affords additional support so that both hands are dedicated controllers, and often means to tilt the device to better accommodate an ergonomic view angle (see figure 5.1).

---

[11] On the iPod Touch, location based services and phone/video calls are offered using WI-FI, via local triangulation and VOIP applications (e.g. Skype), respectively.

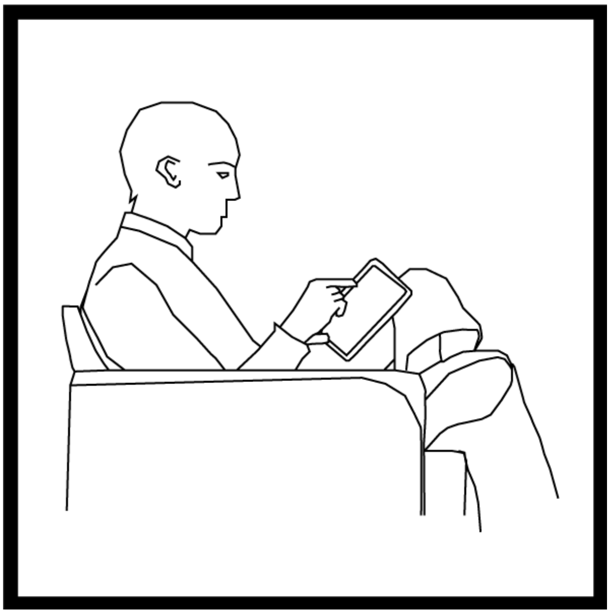

Fig. 5.1: Using a tablet PC while sitting in a chair

3. The hardware interface is comprised of a smooth touch screen surface, with the exception of very few physical buttons.

The smooth touch screen display does not afford our tactile knowledge of physical objects, due to the lack of texture or surface gaps. In terms of tactile affordances, these device are equipped with are close to none, as the smooth surface offer few tactile cues for the user. Different application made available for the devices, seek to emulate textures visually. As an example, the iPad app “KITTY!”, available from the App Store, aims to simulate the stroking and cuddling of a cat. By touching the screen, the kitten starts to purr and imprints of your stroke can be seen in its fur. When you stop stroking the screen, the kitten goes back to sleep.

With the lack of tactile cues, and the positioning of the display and control features on the same surface, the user, albeit mobile, needs to visually confirm his actions by looking at the screen (see figure 5.2). HCI expert Russell Beale points to an interesting fact relating to screen size of smaller portable devices in terms of engaging with the environment: “The smaller screen height [...] also presents less of a social barrier to visual communication with other parties, making them seem somewhat more acceptable for use in meetings and other such settings” (Beale 2009).

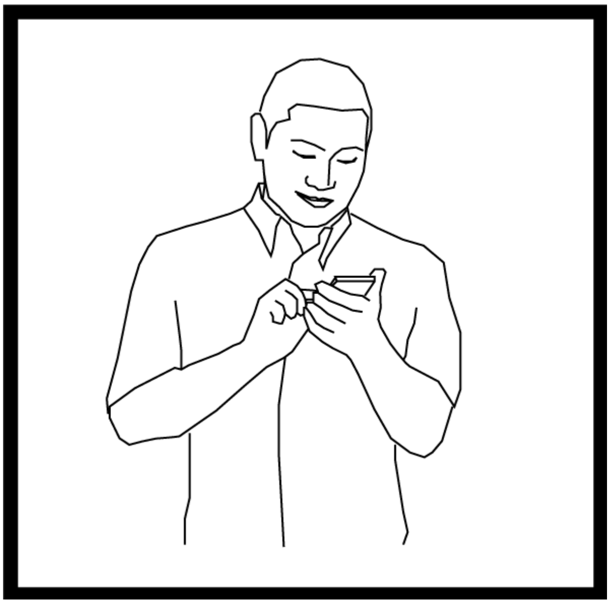

Fig. 5.2: Line of sight and common body-posture when using a smaller mobile devices.

4. The user interacts with the device by looking at the display, and navigating and manipulating the software interface through direct, multi-touch gesture control (pointing, pressing, and sliding fingers over the touch screen surface.) The devices do not accommodate any haptic feedback, and vibration and audio are the main non-visual feedback mechanisms. As mentioned earlier in Chapter 4, there are no common standard for gesture control. Different touch screen devices, as well as different applications available for the device, are interpreting touch-gestures differently. The lack of a common gestural command language leads to inconsistent and sometime confusing use of particular gestures or combinatory gestures, hindering seamless interaction. The gestural interface may be consistent within a single company's product line, as is the case with Apple's iPod Touch, iPhone and iPad. But even in that product line, different third party applications utilize the available touch and gesture controls differently.

5. Third party applications are made available via a pre-installed store application shipped with operating system (e.g the iOS and App Store, the Android OS and Android Marked, and to a lesser extent Window Mobile/Windows7 and the Window Phone Marketplace,) making the devices highly versatile in use, and customizable for the user.

## The mixed-reality potential

### Coupling with sensors

All of the mobile touch screen devices are already equipped with a range of sensors. Haptic input is thus not limited to touch and gesture control. A smart phone today often contains a accelerometer (motion detector), thermometer, light and sound detection sensors, radio wave and wi-fi detectors, in addition to localization services presenting the absolute position, and relative movement in terms of distance or time (based in GPS or local triangulation systems.) The the inbuilt sensor technology affords a range of use.

Increasingly, the device can accommodate position based services. There is a growing field of position-based or marker-based applications, that offer the user access to particular information based on where he is located, utilizing GPS/triangulation functionality, as well as marker recognition software (barcoding, QR-coding, fiducial markers etc.,) built in the hardware device, or made available through applications. Additionally, augmented reality applications were digital information is added as the computer recognizes cues in the environment (QR-codes, bar codes, images, text, or particular physical objects) are becoming common-place. Applications for recognizing bar codes or QR-codes can trigger the retrieval of particular product, event or location-specific information. In other cases, added digital information is layered on top of a real-world objects, by investigating objects or events through the display. One example is Word Lens, an application which translates between English and Spanish in real-time, by displaying the translated words in the same color (and if possible font), on top of the word in the original language.)
Another example is the iPhone application GeoDoodle3D that utilizes the built-in accelerometer, camera, GPS and compass function of iPhone, enabling a user to add a hand-drawn sketch to a specific location and later revisit his drawing through the camera screen. The doodle is added by using the fingers to draw on the the camera screen (the touch screen). The drawing is then coupled with the geographical location and orientation data, and will reappear in the camera window whenever the device swoops over that particular position.

The devices are also highly inter-compatible, as they support a range of standardized methods for sending and receiving data—wireless, infra-red, bluetooth and near field technology—user content and sensor-data can be transferred and synchronized between different computer devices.

The many ways of we can address hardware features and sensors in various applications are truly promising, and we have most likely only scratched the surface of what such devices can do.

## 5.2 Responsive Surfaces

The power of touch screens demonstrated in various mobile devices, have been continued in the development of larger interactive surfaces, optimized for participatory interaction between several users and available content.

### Interactive learning

In the 1960 article “Man-Computer Symbiosis,” Licklider discusses the potential of cooperative interaction between man and machines, and points the importance of collaborative media that allows people to think aloud together.
“Nowhere, to my knowledge, however, is there anything approaching the flexibility and convenience of the pencil and doodle pad or the chalk and blackboard used by men in technical discussion” (Licklider 1960).

The blackboard, and later whiteboard, have played a key role in education and business meetings for decades, precisely because they accommodate so many communicative features. Content can be presented on the fly, by moving a chalk stick or a pen over the surface, and with a single swipe the same content can be erased or slightly altered. And whatever is presented is immediately visually accessible to many, who can partake and add to the content message.

#### Interactive whiteboards

Interactive whiteboards carries many of the tactile affordances of the earlier old blackboard and whiteboards, and bring several new features to the table. Interactive whiteboards are responsive displays, that allow a user to manipulate virtual content projected onto the surface, with a finger, a stylus, or remote. The content is projected onto the IWB surface, is the projected desktop image of a connected computer, turning the IWB to a huge touch screen. With simple gestures content is moved, resized or erased, and media content available on the computer is immediately available for presentation. The IWB surfaces are made touch sensitive, either through the application of infrared scan technology (small IR-cameras scanning the display surface, and register point breaks in the grid,) or resistive touch technology (pin-pointed pressure applied to a surface membrane so that it connects with the conductive back plate.) This separates them from many touch tablet PCs and smart phones based on capacitive touch technology (coated layer on the display surface, that registers electrical charges generated in contact with human skin) utilized in most smart phones and touch tablets.

Due to the choice of touch technology, and resistive touch technology in particular, interactive whiteboards do not support multi-touch use (only single-touch point, click, and slide navigation,) and displays do not work well in sunlight (as the surface layers reflect the sun.) However, resistive touch technology is very precise and affordable, a combination that is part of the reason why it has become so popular in educational tools. A 2011 survey by Education Research Market states that 63,5% of all educators in K-12 schools (primary and secondary schools) in North America, have “a dedicated interactive whiteboard in their classroom” (Selling To Schools 2011).

The IWB hardware interface promotes motor-centric and touch based interaction, allowing the user to stand and manipulate virtual content by directly touching a screen. However, the devices are often configured to work with computers running GUI-like interfaces, and user touch input is usually translated into left and right mouse clicks.

## Tangible Computing—from touch to grasp

The concept of tangible computing or tangible user interfaces was first introduced in the presentation of the Tangible Bits projects, developed by members of the Tangible Media Group of MIT back in 1997. The goal of projects was to create a computational environment that allowed “users to 'grasp & manipulate' bits in the center of users’ attention by coupling the bits with everyday physical objects and architectural surfaces” (Hiroshi Ishii and Brygg Ullmer 1997).  Inspired by the tactile affordances of historical scientific instruments (consisting of buttons, levers, knobs and turning wheels), and the belief that interaction should be based on how man operates his body, and manipulates physical objects, Tangible Bits was presented as interactive environment consisting of several interfaces working together to optimize man-machine interaction. The horizontal metaDesk and the vertical transBOARD, were both interactive surfaces that supported the use of physical icons and specialized instruments as navigation and manipulation devices. And both interfaces were framed within in particular space, the ambientRoom, where soundscapes and lighting could be optimized to fit various tasks. Tangible Bits, as such, never became a standardized computational environment, but the concept of tangible computing and tangible user interfaces are well established. The transBoard has later been pursued in different interactive whiteboard devices, and wall-based multi-touch screen. And hardware interfaces similar to the metaDesk are already available on a commercial market.

### Microsoft Surface

Microsoft Surface is a large (70-100cm) responsive multi-touch display, where content and functionality is accessed by issuing touch-gesture commands. The Surface System was

launched by Mircrosoft in 2008, and was designed to accommodate four criteria of interaction:
First, interaction is to be direct, a user should be able to *directly manipulate* any content, without the use of input or control device. Secondly, the system should accommodate *multi-touch* interaction (per today a Surface system can keep track of up to 50 individual touch points). Thirdly, the system should support *multi-user* interaction, mainly made possible through its multi-touch support. And finally, the system should *recognize the presence and orientation of object* on the surface (Fletcher 2009).

The multi-touch features turns the often private display into a collaborative surface, where many participants can investigate and manipulate content simultaneously, by directly touching and manipulating virtual objects. But in my opinion, the most important feature of Surface systems are their ability to recognize tagged physical objects. The object recognition feature of Surface is based on a technology called Pixelsense (Microsoft 2011) that generates images of the object on the surface. The processed images are interpreted by image-recognition software that identifies an object from a predefined list of tagged objects.

Surface systems are commercialized, but with a price tag starting at $12,000, the average computer user does not have the option of incorporating such a system as part of their computer environment. The Surface device in itself allows for varied use (the software platform for the device is the same as one utilized in Window 7 Touch), and Microsoft has already released a beta version of a software developer kit, (with an official one about to be released), to enable developers with a Surface system to create new software applications for it. However, currently the software interface and applications are particular to the systems' early adopters. Surface is promoted to the financial, military and healthcare sector as a collaborative and interactive planning and simulation tool, to retails stores and product conferences as a shopping experience enhancer (as costumers and spectators are invited to place the physical products on the surface to learn more about them,) and to hotels, restaurants and bars, as an entertainment device—giving their customers access to an easy to use multi-user gaming platform. And for now that, given the price and range of applications, the Surface environment is a user interface for the fewer of us.

### The power of phicons

The object recognition feature of the Microsoft Surface supports the notion of seamless interaction between physical objects and virtual content. Recognizing objects based on a processed image is one thing—having physical objects that can display and transfer content between themselves and a host computer is something much more. And these objects do exist.

Phicons, or physical icons, are small, cube-shaped microcomputers (CPU-based) equipped with tiny displays, wi-fi radio, movement sensors (accelerometer), and near-field communication technology (or alternately, infrared transceivers). This means the cubes register when they are adjacent to another cube, when they are moved, or when they are shaken or tilted.

Within the Tangible Bits environment, the physical icons played a key role in accessing and navigating virtual content, but the technology they were based on did not catch on. Siftables was prototyped by David Merrill and Jeevan Kalanithi at MIT in 2007, and were "inspired by observing the skill that humans have at sifting, sorting, and otherwise manipulating large numbers of small physical objects" (Merrill, Kalanithi and Meas 2007). Later Merrill and Kalanithi founded Sifteo, and commercialized Siftables as Sifteo Cubes—inexpensive game devices, instantiated by game software running on a host computer. However, the technology employed presents a much broader use, in terms of creating tools and applications that fronts moving and grouping of digital information, as we would move and group physical objects. And Sifteo has released software developer kits inviting people to do just that. At MIX11 (an annual technology conference, hosted by Microsoft, gathering developers and designers working on high-end web solutions), the experts connected to the Microsoft Surface team demonstrated the Siftables in conjunction with Microsoft Surface. The demo featured a task assignment application, where images of people where presented on the display were dragged and dropped on a Siftio cube. Each cube was programmed to hold a specific task. By holding the different cubes next to one another, performing a certain shake movement, groups of people could be connected to same task, while positioning the cubes differently resulted in tasks being redistributed between the group members (Barraza et al. 2011).

A similar use of phicons and surface computing is demonstrated in the Reactable platform developed by Martin Kalkenbrunner. Reactable is a multi-user touch screen synthesizer, where users create sound patterns and songs, by adding and moving phicons around on the table (see figure 5.3). Each phicon represents a sound, effect or modulation, as well as acting as synthesizer control devices. Proximity and distance between these objects shapes the sound scape, tempo and volume. Learning what the different phicons do, in terms of being able to predict how sounds are trigger and adjusted take some time to learn (and some inherent knowledge about digital sound processing,) but operating the interface seems intuitive to all. There is direct correlation between the placement and movement of the objects and a corresponding change in the soundscape.

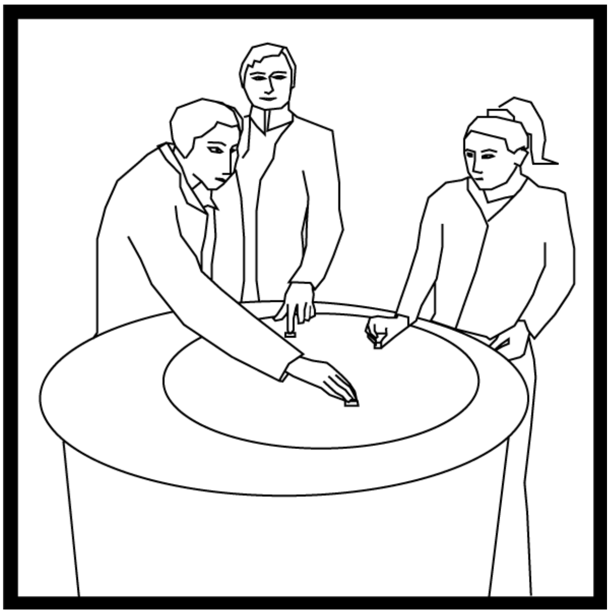

Fig. 5.3: Several operate a responsive surface with phicons.

Responsive surfaces have the benefits of the touchscreen. With the aid of fingers and hand gestures, users can directly manipulate and navigate through digital information. The increased display size invites participative interaction, as the surface is visible to a larger audience. Single-touch surfaces are good for holding interactive presentations, while surfaces accommodating multi-touch allow several users to simultaneously interact with virtual content.

Due to size of the displays, these computational environments are not particularly mobile. IWB are in danger of becoming an oversized desktop computer, where the interaction mode equals to the point and click navigation of the GUI. The Microsoft Surface device can equally easily be equipped with the GUI-based Window 7 Touch interface. So even though hardware interfaces are optimized for haptic interaction, the software framework might be framed in a graphical user interface, promoting a WIMP-like interaction style, where touch-gestures are simply translated to points and clicks.

Although in its early stages, the coupling of phicons with surface technology allowing the user to physically transfer information between displays and computers—without having to consult a GUI to copy and paste data between local or networked folders—is truly suggestive of the potential of coming user interfaces. As applications made for Microsoft Surface using the software developer kits are partly applicable on screens supporting Windows 7 Touch technology, the price of obtaining one version of such a system is becoming affordable.

## 5.3 Responsive Environments

Responsive environments literately place the body in the center of the interface. Components consist of motion-tracking systems that pinpoint a user's position and movement within a defined, limited space. Users can issue system commands, and manipulate virtual object, either with the use of hand-held devices (Wii-Remote or similar WI-FI or Bluetooth based controllers)—or by gesture and face recognition software capable of translating gestures into commands.

### From Virtual Reality to CAVEs

Virtual Reality did not come to pass partly because the devices employed in the interaction, the huge head-mounted displays, encumbered its users. But more importantly, main assumption of the virtual reality paradigm, namely that full computer immersion in virtual environments could be achieved by stimulating the visual sense only, just did not hold. However, VR technology did not vanish, rather it has been employed in creating environments that accommodates more, or all of the human body, and car and flight simulators are current examples of this.

One implementation of VR technology that has been particularly successful is CAVEs, where the virtual environment is projected, or mapped onto, a physical three-dimensional space. A CAVE (Cave Automatic Virtual Environment) is an interactive room, where three to six of the walls are presenting virtual information to the user. The virtual environment is created based on the user's position, his movement and hand gestures. It is a surround-screen system, where the body of the operator is what actuates interaction.

The first CAVE was finished in early 1992 by the Electronic Visualization Laboratory (University of Illinois, Chicago,) and was presented as a virtual reality theater (Cruz-Neira, Sandin, DeFanti, Kenyon and Hart 1992). The environment consisted of 4 interactive screens (three walls plus the floor,) and a surround sound system. The system was presented with a range of rather specialized programs optimized for the CAVE, among them a three-dimensional weather system map, a brain-surgery planning software, a program for exploring the evolving universe, and a molecule modeling software. The first CAVE was thus very specialized in use, but demonstrated an stunning ability to visualize and scale the very small or the very big, to human size, allowing the user to be immersed in, exploring and manipulating, micro and macro worlds previously removed from us.

The CAVE hardware interface comes in many versions and sizes, but each of them share a set of features. The illusion of immersion is generated by projecting stereoscopic computer images (high definition 3D) on the walls, where mapping and rendering software fits the video images to the exact dimensions of the space. The user employs 3D glasses to correctly interpret the layered image information, thus extending the notion of depth beyond the physical CAVE walls.

Through the use of hand gestures or hand-held control devices, the user can navigate through, and activate features presented within the virtual environment (see figure 5.4).

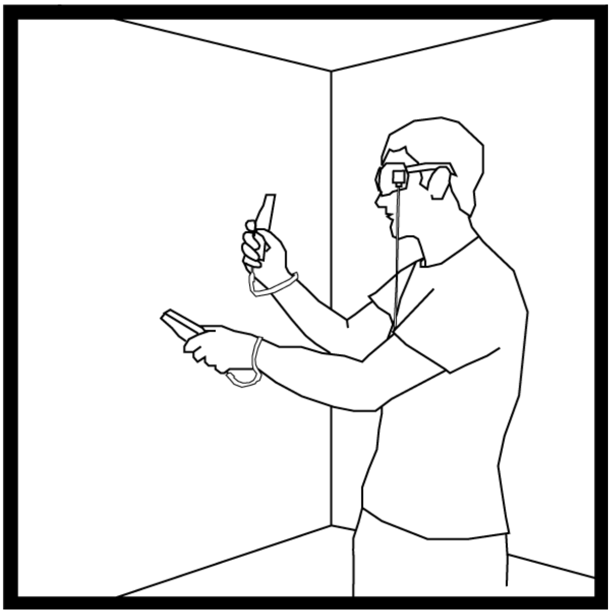

Fig. 5.4: Interacting in a CAVE environment.

As the 3D glasses are semi-transparent, the user can see himself in the interface, allowing the real to blend with the virtual. The user is mobile, he can walk within the CAVE, and his position is continuously tracked to project the a correct perspective to that of the users point of view. Several CAVEs are also equipped with a sound system, and utilize auditory feedback as part of the interaction. The first CAVE only supported single-user tracking and navigation, but later editions accommodate multi-use—although that usually means that one main user controls the viewpoint, while the other participants are enjoying the ride.

Programming and preparing a CAVE for a particular application requires an elaborate mapping and configuration process, where each scene needs to be rendered to fit the particular responsive environment, thus making it less versatile or general in it use. However, they are becoming important visualization tools within scientific research, product development, and military training, and many CAVE systems are thus found within universities, and research and development departments of manufacturing companies. The immersive qualities of CAVE systems are also explored in interactive art and games, where the artist has a unique opportunity to let the spectator  become an exploring participant within the virtual environment. Artist and associate professor at the H.R Hope School of Fine Art at the University of Illinois, Margaret Dolinsky, has produced several works for CAVE systems, and has a particular emphasis on the participative and collaborative art experiences. In her early work *Dream Grrrls*, developed by Dolinsky and programmer Grit Sehmish back in 1997 for the original CAVE system (Electronic Vizualization Library 2011), the participant is invited into a dream world populated by characters previously only presented through Dolinsky's oil paintings. Dolinsky seeks to present a dream world to a fully awake participant. Since the environment is presented from the users point of view, the act of sitting down or moving will alter the perspective of the environment, as we would expect in the real world. And as with dreams, the act of walking down an illuminated path, is no guarantee for it suddenly turning into a portal, that brings you to new and unexpected places. Trigger other cues, and gravity is no longer a consistent force, as the participants views himself being elevated through the virtual environment. *Dream Grrrls* extend and augments user perception, and allow us to do so by moving within the CAVE space, and navigating and manipulating virtual objects with a hand-held device.

The last decade has produced fewer pure CAVE interactive art works, partly because the elaborate process of image mapping and rendering. But I also reckon, that due to the advancement within 3D modeling in the game industry, there is a demand for more realistic looking 3D virtual objects, to have satisfying immersive experiences.

And some evidence of the what the prominent use future CAVE systems will be, might be shed by the announcement of the Next Generation CAVE, also conceptualized by the Electronic Visualization Laboratory and funded by the Computer and Network Programs of the National Science Foundation in the U.S:  “The Next-Generation CAVE Virtual Environment, or NG-CAVE, is a scientific instrument that enables researchers to visualize data in a fully immersive 3D stereoscopic environment; it serves as the lense of a 'telescope' or 'microscope,' enabling them to see their e-science datasets that reside in cyberspace.” (Electronic Visualization Laboratory 2009)

The CAVE systems do support several significant, and yet very specialized operations, due to the cost and physical space involved in setting up the system, and development of applications. With the advent of Microsoft's Kinect in early 2009, a new kind of responsive environment were made available to the general man.

## From Artificial Reality to Kinect

In chapter 3, I introducted Myron Krueger's concept of Artificial Reality, and its implementation in the VIDEOPLACE installation. Krueger sought to remove all hand-held controllers navigating and manipulating the interactive environment, as they distanced the user from the interaction, separating the human from the machine. Instead he proposed using the human body, its ability to move, and an intuitive gestural sign language as the means of interacting. Ideally, the human body should be the interface, and the mode of interaction is operating the body as one would in everyday interaction with people and objects. This is the key to immersive man-machine interactions.

The VIDEOPLACE installation was set in a located and limited space, and consisted of several projectors, a video recording system and specialized video processing software. The video recording system recorded the user's posture, positions, movements and hand gestures within the defined area. The video material was then processed by a specialized software digitizing and flattening video of any user actions to a two-dimensional silhouette, which is what is represented via the project system and how the user is mirrored in the system (see figure 5.5). The posture, position, movement and gestures of the silhouette that manipulated virtual objects or trigger events within the installation, and the silhouette is continuously re-activated by user's actions. All actions and projections are happening in real-time. There was close to no lag between issuing a command or triggering an event, and the projected summary of the screen—connecting the spectator with his own image. The real-time presentation of the system was key to Krueger's thinking about interaction. No lag meant that the user was in immediate contact with any of the virtual objects or events presented within the environment, merging, or creating a seamless coupling between the virtual and the real.

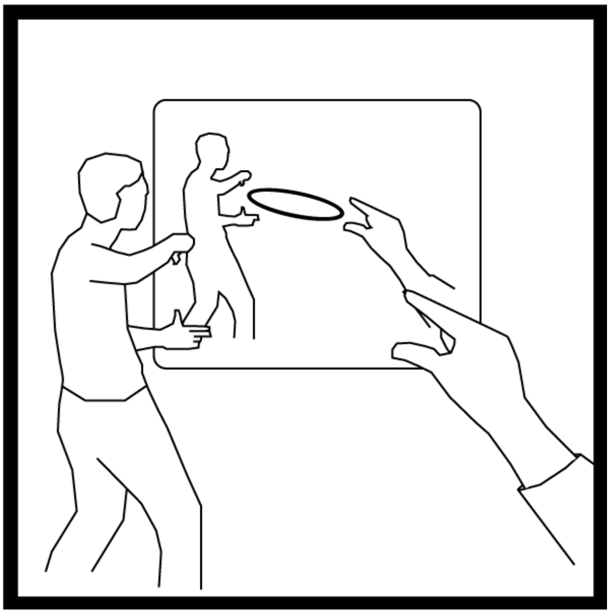

Fig. 5.5: Interacting within VIDEOPLACE.

VIDEOPLACE allowed the user to see himself while interacting with a virtual object in real-time, and while some of the installations programs or instances, presented the onscreen silhouette according to actual user movements or gestures, others deliberately distorted them. The responsive environment also allowed for several users to participate at the same time. Several users actions could be recorded from different locations, while the projection system output the same merged summary of all of the user action to all of the participants. Kruger thus thought that VIDEOPLACE like interfaces were particularly optimized for telecommunication applications, as it instantly allowed user to be present in the same environment (albeit virtual,) performing actions in real-time. The VIDEOPLACE programs made the participants see and become aware of responses, originating in their proprioceptic sense—or as Andy Cameron stated in a discussing of Krueger's work—interaction within the VIDEOSPACE environment is about “making the body proprioceptively aware of itself” (Cameron 2009), as it triggers powerful mirror-effects where the full-size body is involved. Over 25 programs were made for the VIDEOPLACE installation ranging from drawing applications, simple games involving virtual objects (playing with a graphical string) or virtual animals (as in Critter,) or mirror-distorted compositions (as is the case with Replay.)

Myron Krueger stopped the work with VIDEOPLACE in 1984, and the environment was permanently installed at the Connecticut State Museum of Natural History (and presented as part of the museum program until early 2000s.)
Another 10 years had to pass before a proper alternative to VIDEOPLACE came about. It was not an art installation or stimulator, but rather a game controller.

## Kinect—the power of computer vision

Launched in 2010, Kinect was Microsoft's top of the shelf controller for the game console Xbox 360. It separated itself from all other controllers as it did not require the gamer to use any hand-held device or touch any controller. Instead the Kinect tracked the user movements and hand gestures within a limited space, and translated those into game commands, allowing the gamer to operate the game environment with his body. The Kinect hardware consists of a depth sensor, a RGB-camera, microphones (allowing for voice commands), which fuels the inherent face and motion recognition software.  The mapping of user actions, are represented to user via a display. As with the VIDEOPLACE interaction, the user see himself within the virtual world manipulating virtual objects in real-time. The user can be presented as himself or as an avatar that operate according to users movements and position.
The first reviews of the Kinect were divided. As a game controller it was pretty expensive, close to matching them price of the console itself. In addition, an optimized Kinect environment required a fitting physical space, and while it, initially, was exciting to issue the different gestural commands early on in a game, for many gamers, they turned out to be more straining than fun after hours of play. That was the game review...

The hardware and software components of the Kinect are closed-source (i.e. proprietary,) but within days of its release; developers, artists and engineers had come up with a staggering number of alternative suggestions for how the device could be employed. And the competition was on. A competition in understanding the current limits of the device, while seeking ways to access the data collected and processsed by the device. And within weeks Youtube.com and Vimeo.com bloomed with Kinect hacks, or to be precise, Kinect applications that made use of open sources drivers for reading the data passing the USB connection of the device. The launch of open source drives ensured open access to any sensor data collected by the device, and a whole group of third party developers started building applications, programs, art installations and learning environments based on their unique interpretation and presentation of sense data.

### Interactive art

Kinect was welcomed in the electronic art world as soon as it was released, and with the launch of open source drives it became a viable instrument for instantiating interactive experiences. In terms of creating participatory interactive environments, Kinect is an affordable, mobile device that simplifies and extends the potential of current mapping and tracking software, and video camera recording systems.

The most common Kinect setup till now has been to incorporate user actions within a virtual environment and displaying the summary of the mapped and the virtual via a display. So the system allows for a haptic exploration and manipulation of the environment, and results of user actions are presented via visual feedback. But Kinect's ability to map and track user movements has also been applied in other interactions.

The Spanish artist collective BlablaLab skipped the display altogether with the urban intervention project “Be Your Own Souvenir” (BlablaLab 2011). The project was first presented at La Rambla in Barcelona in early 2011, where the collective invited people into strike a pose on a specific location of the street. The area was encircled by several Kinects, mapping the full-size body of the poser from all angles—taking a 360 degree posture shot of the participant. This three-dimensional snapshot was then processed by a 3D printer, producing a miniature human sculpture of the participant on the fly. Kinect has also proved to be a powerful controller of physical drones. With simple hand and arm gestures ETH Zurich researchers have demonstrated how Kinect can be used to fly a quadrocopter (Smalley 2011.) The last example is of particular interest as it demonstrates Kinect's power as interface control device, where physical objects are manipulated and controlled, without users having to interact with a display.

### Interactive education environments

One year after its release, several academic papers have surfaces discussing the potential of Kinect. Hui-mei Justina Hsu from the Fo Guang University, department of Learning and Digital Technology has paid particular attention to the use of Kinect in learning environments and states:

> “First, Kinect is a flexible teaching tool. Teachers can interact with contents via body movements, gesture and voice without using keyboards or mice. Second, Kinect can accommodate multiple users; therefore, students can have a fare share of control over interactions. A Kinect-enabled classroom can support whole-class instruction, group work and teacher-student one-to-one interaction. Third, it is a versatile tool. As it collects 3D information, Kinect can support various teaching activities such as dance and martial arts. Special instructional design can be implemented to reinforce the connection between teaching contents and student physical responses”(Hui-mei Justina Hsu 2011)

Hsu believes Kinect to hold the affordances needed to boost student motivation, to participate and interact with the subject matter. Kinect is a key component in creating a range of interactive environments exploring particular themes, that we can operate using a natural controller—our bodies. As it also accommodates participatory experiences, students have the opportunity to interact with content knowledge either on their own, in groups or with their teacher.

Well-known interactive artist and media professor Golan Levin is currently using Kinect and Computer Vision technology as a starting point for teaching his students interactive art and computational design. Students are encourage to embed Kinect functionality to upgrade their computational environments, and within the course periode of a few months, students had produced several augmented reality games, specialized gesture-based controllers for VJ (video jockey) and game environments (Golan Courses 2011).

The use of Kinect, or Kinect-like devices is still in its infancy. We have only scratched the surface in understanding what such devices offer, in terms of creating valuable interactive entertainment systems, art experiences, and learning environments.
Kinect on its own it not sufficient to create an interactive environment, it demands complimentary hardware and software to be operational. Nonetheless, the functionality of the Kinect is of such a general quality, that when well staged, presents itself as an invaluable component in any conceptual model of interaction.

## 5.4 Wearable Computing

Wearable computers points to a class of smaller electronic devices that are attached to the user's body. The devices are equipped with various sensor technology—to measure position, movement, temperature, pulse— and the computational brain is embedded within or connected to the device via a mobile host computer. Many wearable devices are intended for specialized use. By coupling a GPS with jogging shoes, you will have access to where, when and how fast your moving. But what all of the devices have in common is that they all collect and process data from the vantage point of the wearer's body.
There are many specialized wearable computing devices, however, I will take a closer look at a one particular device that has the promise of becoming rather versatile.

### Sixth Sense

The Sixth Sense is a conglomerate device—a wearable gestural computer, consisting of a web camera and a projection system both connected to a mobile computer (e.g. in the form of a smart phone,) a mirror, and a gesture-based 3 dimensional mouse. The camera, projection system, and the mirror hangs from the user's neck, while the mobile computing device is placed in a trouser or sweater pocket. The device is completely portable, as it depends on no infrastructure.

Color markers[12] attached to both thumbs and index fingers, turns the four fingers into a three-dimensional mouse, where a wide set of gestures are recognized and translated into computer commands—all based on the position and movement of a single color marker relative to another. The camera can, with the aid of image recognition and marker technology, capture and recognize a range of physical objects, as well as tracking hand gestures. The projection system is battery powered, and can project on any surface. And as the device is mounted around the neck of the user, a mirror is added to the setup adjusting the angle of the projection to ergonomically fit the position and posture of the user. Microphone and headphone are optional features. See *figure* 5.6.

Sixth Sense technology proposes a mode of interaction where the physical world is augmented and extended with digital and sensor technology. The user is completely mobile, and is only limited by battery power, and connectivity (using applications depending on access to WI-FI or mobile networks.) Any surface can be turned into an interactive space, where the camera projects the application environment of choice onto the surface, and the user navigates and manipulates digital content with hand gestures. As the software platform consists of image recognition and marker technology, the camera can recognize physical

---

[12] In the current Sixth Sense prototype, the color markers are history, and gestures are recognized by depth-sensor technology similar to that demonstrated in the Kinect devices.

objects, and add relevant information about it, by projecting it onto the object or a nearby surface.

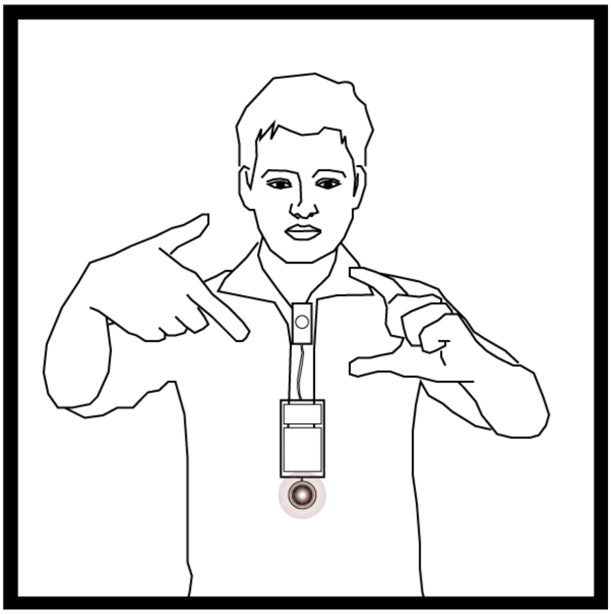

Fig. 5.5: Wearing and interacting with a Sixth Sense device.

The Sixth Sense technology (surface computing modules and image recognition software) was developed by PhD student Pranav Minstry, part of the Fluid Interfaces Group of MIT Media Lab, and is based on the *Telepointer* prototype developed by MIT student Steve Mann in the late 1990s (Steve Mann 2000). A fully working prototype of the Sixth Sense device was showcased at two TED talks in 2009, by Fluid Interfaces Group founder, Pattie Maes, and Pranav Minstry, respectively[13]. The showcase featured the Sixth Sense with its current software applications: Telephony (a number pad is projected on any surface, allowing the user to dial using his fingers), navigate maps (zooming, adding locations), clock (projected on any surface), augmented reading (identifying books from their title and by searching various knowledge bases, and adding info about book author, where they can be purchased, and relevant events,) drawing applications, zooming features, taking pictures (using the hands to frame the object,) as well as other augmented reality features— where text strings, images or barcode/QR-codes markers, acts as identifiers for relevant digital media content that can be projected onto or next to the physical object.

The device was an immediate media sensation, and the developer team won several inventor awards the same year. Sixth Sense technology combined the functionality of a mobile smart phone, surface computing, and gesture and image recognition features found in devices like Kinect. The hardware components constituting the device were very affordable (the current

---

[13] Pattie Maes' demonstration of Sixth Sense at TEDTalks, February 2009, and Pranav Minstry's demonstration of Sixth Sense at TEDIndia Talks November 2009, can be found at http://www.ted.com/talks

prototype costs about $ 350), and all components were easy to get hold of. Most promising was probably the announcement that Sixth Sense technology should be open source. Then from late 2009 to early 2011 not much more was heard from Minstry. Until he was interviewed in late February 2011, by journalist Jesse Brown from Canada's weekly current affair magazine MacLean's (Brown 2011). What surfaced during the interview was that part of the Sixth Sense operating system and overall software platform was based on proprietary code from Microsoft, meaning that the developer team needed to rewrite the code for the operating system, and relevant software running on it—to be able to present Sixth Sense as an open source project. The Sixth Sense core developer team has now launched an open hardware and open source software project on Google Code, inviting people to build their own Sixth Sense device, and develop new applications for it (Sixth Sense Google Code 2011). However, the site still lacks relevant software documentation and code, for this kind of activity to start.

The Sixth Sense technology may revolutionize general-purpose computing, if enough productivity applications are developed for it. It supports truly uninhibited surface computing —as *any* surface—be it your hand or a building wall, can be activated with a single gesture. Other key features of this computer, is the secondary role of the display. The computer sees the world from the user's perspective, and the user communicates with the machine through haptic input.

## 5.5 Case Conclusions

### The portable, affordable data transceiver

Mobile touch screen devices becoming increasingly common, and are used very differently from owner to owner. The multimodal qualities of the device, can be measured in the of haptic input they favor. The devices promote touch and gesture control, as well as acting on proprioceptic response, due to inbuilt sensors, such as accelerometers, and via GPS technology. The lack of tactile affordances, such as texture or surface gaps, forces the user to address the display in terms accessing and navigating device functionality and content. The size of the display thus plays a significant role in terms of what operations the device affords. While it is good for shorter sessions of accessing media content, checking emails, or playing games—the size of the screen restricts the amount of information that can be displayed at the same time, making it less attractive as a productivity tool, a restriction furthered by the fact that display doubles as an input control surface.

A truly promising feature of the devices lies in coupling its with sensor technology, as wells open APIs inviting third party developers envisioning further uses of the devices. The devices are also highly inter-compatible, due to wireless, infra-red, bluetooth and near field technology, allowing for transfer and synchronization of content and meta-data collected by its sensors, between a range of computer devices.

Mobile touch screen devices are promoting a wide range of use, due to inherent sensor technology and stunning span of software applications. This, combined with being highly portable and inter-connected, make them versatile general purpose devices on their own, even though their size and form factor discourages longer, focused working sessions.
As a complementary device in a larger computational environment, they are proving to be significant.

## The shared plan desk

Responsive surfaces, such as the Surface system or Reactable do, as mobile multi-touch devices, allow for haptic input via touch and gesture based control. And as these devices, responsive surfaces position the display and control features on the same surface. However, the surface display is so large that they provide users with a good overview of content and area for navigation. The size of the surface and extensive multi-touch support, promotes participatory interaction, as many users can operate the computer at the same time. They thus encourage visual brain-storming and planning sessions, as wells as sorting through and arranging larger volumes of information. The mere size of such a computational environment, fixes it a to particular location.

The object recognition feature of the surface technology, has a lot of potential, in terms of letting physical objects initiate access to related information. As the position of a physical object dictates the initial orientation of information invoked by it, several users can simultaneous work within the interface from their individual point-of-view. The responsive surface is smooth, non-texturized, and does not provide the users with any tactile cues in itself. But combining surface technology with the use of phicons, virtual objects and data clusters can be manipulated as we would a physical object—by dragging it over the surface, picking it up, or moving it to another location. At the same time, phicons are not restricted to the limitations of individual physical objects, as they allow us to group, merge, and split the information they carry with equal ease, with simple hand gestures and movements.

Surface technology is still in its early stages of development, and is currently supported in rather expensive devices.  Equally there are relatively few applications developed for it as of now, to call it a general-purpose device. The size of the display has the promise of allowing user to visually present a vast amount of media information at once, providing him with an overview and the ability manipulate this content directly. Phicon-technology is promising in itself, introducing physical information carriers that transmit and receive data, with touch and proprioceptic gestures known from our everyday interaction with objects.

## Proprioceptic control and mirroring

Responsive environments are the most promising computational environments promoting proprioceptive control. Body movement, position, and hand gestures are used to issue commands and navigate the environment, providing a real-time interactive experience and manipulative control of virtual content.

CAVE applications are particularly interesting as user is surrounded by a three-dimensional screen, that continuously present visual information in several dimension, mapped to his point of view. They place the user *within* a virtual world, while allowing him to explore the world with his full body. Feedback is primarily visual (and often auditory,) and powerful as both the user's direct gaze as well as peripheral vision is acknowledged. And some CAVE implementations, vehicle simulators in particular, are equipped with haptic force-feedback systems, giving user a relevant proprioceptic response to their actions.

In the early VIDEOPLACE installations and later CAVE art installations, proprioceptic input is mirrored in the surrounding display in real-time, either directly (symmetrically,) or purposely indirectly (distorted, delayed or multiplied)—letting the participant become immediately aware of the effect of his body. It is within such responsive environments that the significance of Hansen's argument becomes apparent. These environments have a unique power to extend the body schema, as they allows us to engage with our motor sense in completely new ways.

Kinect, being a portable depth and motion detector that can be addressed from software running a regular desktop computer, adds the promise of easily moving responsive environments between locations. The employment of Kinect to directly translate body movements into orientation directions for physical drones, obliterates the need for any display. The human body has become the interface, and the machine is solely, responding to proprioceptive control.

Responsive environments are significant multimodal environments, capable of introducing qualities of the virtual in direct and meaningful ways. They can induce powerful art experiences and produce rich learning experiences, precisely because the both the visual and somesthetic senses are employed to interact with the virtual.

### Mobile surface computing

Sixth Sense technology is conglomerate, bringing in the affordances and functionality known from mobile phone devices, surface computers, and gesture and motion detectors.
The fact that the device can be operated without an external infrastructure, and its ability of turning any surface into a temporary computational environment, are constituents for making it a truly portable computational device. Equally important is the removal of a physical display. Rather than accessing computer functionality through a display, hand gestures invoke different applications. Certainly enough, in several applications the projected surfaces act as virtual screens the user will have to navigate in terms of achieving certain tasks.

The device itself is currently only prototyped and few software applications exist for it as of now, but the system itself a vivid example of thinking-outside-the-box in terms of framing general-purpose technology. A computer is basically a versatile system where a computational process is modulated through connecting sensors (detectors) and actuators (input devices, controllers, feedback mechanisms,) and can come in any number and forms, depending on how we imagine it. The Sixth Sense device demonstrates just that.

## 5.6 The Promising Future of Personal Computing

What these cases demonstrate, are that the components for building general-purpose computational environments that consider and utilize the touch and proprioceptic senses, to ease the interaction between man and computer, are in place. Haptic hardware interface components are made available commercially, and decreasing in price. On the software side, Software Developer Kits (SDK) or API's are increasing made available, and open source projects for developing program applications for particular hardware devices are initiated.
All of these efforts enable new software solutions for different devices. The multitude of software applications for smart phones and touch tablet PC, made available in a short time due to openly accessible APIs, is a key example of that.

Several of the cases above illustrate that even though the hardware interface accommodate haptic interaction and direct manipulation of virtual content, the software environment and applications may not. It is fully possible to add present a Sixth Sense device or a surface system with a graphical user interface. All of the interfaces described above have elements of a natural user interface. Interaction is based on how we operate our bodies, existing touch and proprioceptic knowledge about what objects can do and how they can be manipulated—so that interacting with the device or within the environment is intuitive.

## Ideal for me is ideal for you?

The general-purpose computer was framed within a workplace metaphor from the late '70s to the late '90s, until the introduction of the World Wide Web and mobile laptop systems, led to an increase in when and where people were using computers, as well as significant increase in types of actions they performed on them. After the turn of the millennium, the workplace metaphor is no longer suitable for describing the interaction mode of the average user. Instead, the personal computer achieved a ubiquitous quality similar to common household appliances, suggesting a metaphor based on that of the home, to describe the position of the computer to the average user. A range of software applications, and internet services extended the computer to become our communication channel, media library and entertainment console, in addition to being a versatile production unit. But even as the the range of human activity administered through the computer increased, the overall hardware design for the personal computer remained static in its form and setup. And it has become increasingly apparent that GUI/WIMP setup framing the general-purpose personal computer, is insufficient for successfully handling such versatile use. The personal aspect of computing should be dealt with in its totality. Users want to do different things with a computer, at different times, in different places, and the computational environment of a single user should reflects those needs. Perhaps the most suitable metaphor for describing personal computing to day should be “me”. The ingredients for building diverse computational environments are certainly in place.

The mobile touch screen devices containing a range of sensor technology, robust screens enabled for surface computing, physical icons, portable motion and depth sensors, and easy-to-use object and gesture recognition software—can all be considered components for building individual multimodal computational environments, customized to meet a single user's needs, a conglomerate of interfaces seamlessly integrated with the everyday environment.

# Conclusion

This thesis has explored the rich landscape of design decisions that have produced our current technological environment of personal computers, and what is missing from that landscape. It has sought to update the reader on what attempts are in the making, and the potential of the user in shaping and influencing future decisions of what such personal computing environments can or even should contain, and facilitate.

There are too few discussions relating to what a personal computing environment should consist of, on the hardware side of things. Most discussions are set on how we should interact within a current computing environment. Our choices are limited to what operating systems, browsers, and social media integration/disintegration we want within the current setup. With exception of the ongoing growth in versatile use of mobile phone and media devices, most of us still let the desktop metaphor framed within a Graphical User Interface dictate the possibilities and limitations of day-to-day computing.

The power of general purpose technology in terms of personal computing, has primarily been explored through a proposal for a general purpose interface, the GUI. But this interface is not optimal for accommodating the range of human activity, partly because it channels these activities through a very specific representation, and mostly because it primarily engages the cognitive faculty of the user, ignoring the human body proper.

Interacting with computers on a day-to-day basis, is a significant part of the life experience for a growing number of the world population, and that interaction mode should increasingly address and incorporate the tactile and motile capabilities of our bodies, to extend our experiences.

We have choices if we are willing to unlearn and become computer literates. Just as we see the significance of learning how to read (and not necessarily write books), we should know what a computer is, and what it can do.  We do not need to become programmers, interaction designers or hardware experts, but we should now a computers parts and their individual use and potential. And be aware that the default setting presented for personal computing, is merely that—one suggestion out of many possible configurations. If we can move beyond current ideas of what a computer is, re-invent and retell the stories of what living with a computer means—we as users, can can be active participants in the ongoing technogenesis.

Artists are significant resources in revealing the potential of technology to users.
They are experimental, not bound to criteria of marked or current trends of functionality.
They are story tellers.

Stelarc chips in: “What's interesting about art is that there is a willingness to mess with new media. To entertain the accident. To be enamored by the ambivalent and the uncertain. To allow for the slippage that occurs between intention and actuality. To undermine and expose new technologies. And to appropriate and morph systems into new operational and aesthetic possibilities” (Aceti 2011).

Everyone should know what a computer is and what it can do, and most importantly, what it can do for you. It is not a neutral tool nor an accessory. It is a part of us, a part of you. And you should make it yours. Be the Cyborg you want to be.

# Bibliography

**BOOKS**

Brouwer, Joke, Sandra Fauconnier, Arjen Mulder, Anne Nigten, editors. 2005. *ARt&D : research and development in art*. Rotterdam: NAI Publishers.

Dourish, Paul. 2001. *Where the action is: the foundations of embodied interaction*. Cambridge, Mass.: MIT Press.
Gibson, James J. 1986. *The ecological approach to visual perception*. Hillsdale, N.J. : Lawrence Erlbaum

Egan, Greg. 2003. *Schild's Ladder.* London,U.K: Orion Publishing

Grunwald, Martin, editor. 2008. *Human Haptic Perception: Basics and Applications*. Basel: Birkhäuser Basel.

Hansen, Mark B. N. 2006. *Bodies in code : interfaces with new media.* New York: Routledge.

Hayles, Katherine. 1999. *How we became posthuman: virtual bodies in cybernetics, literature, and informatics.* Chicago, Ill.: University of Chicago Press

Jütte, Robert. 2005. *A history of the senses : from antiquity to cyberspace.* Translated by James Lynn, Cambridge : Polity Press.

Krueger, Myron.W. 1993. *An easy entry artificial reality. Virtual reality: applications and explorations*. Boston: Academic Press Professional.

Krueger, Myron. 1991. *Artificial Reality 2* (2nd Edition). Reading, Mass.: Addison-Wesley.

Laurel, Brenda. 1993. *Computers as theatre.* Reading, Mass.: Addison-Wesley.

Laurel, Brenda, S. Joy Mountford, editor. 1990. *The Art of human-computer interface design*. Reading, Mass.: Addison-Wesley.

McLuhan, Marshall. 2001. *Understanding media: the extensions of a man.* London: Routledge.

Merleau-Ponty, Maurice. 2000. *Øyet og Ånden*. Translated and epilogue by Mikkel B. Tin. Oslo: Pax forlag.

Merleau-Ponty, Maurice. 2004. *The World of Perception*. Translated by Oliver Davis. London : Routledge.

Merleau-Ponty, Maurice. 2002. *Phenomenology of perception*. Translated by Colin Smith. London : Routledge.

Norman, Donald A. 1988. *The psychology of everyday things*. New York : Basic Books

Wardrip-Fruin, Noah, Nick Montfort, editors. 2003. *The New Media Reader*. Cambridge, Mass. : MIT Press.

**ARTICLES**

Aceti, Lanfranco. 2011. "Inverse embodiment: An interview with Stelarc." *MISH MASH, Leonardo Electronic Almanac*, Volume 17, issue 1, August. http://www.leoalmanac.org/index.php/lea/entry/inverse_embodiment/)

Akass, Clive. 2001. "The Men Who Really Invented The Mouse." *Personal Computer World.* November, pp. 24–25.

Beale, Russell and Andrew Wood. 1994. "Agent-Based Interaction." *People and Computers IX: Proceedings of HCI'94*, 239–245.

Beale, Russell. 2009. "What does Mobile Mean?" *International Journal of Mobile Human Computer Interaction*, Vol. 1, Issue 3. DOI: 10.4018/jmhci.2009070101

Bush, Vannevar. 1945. "As We May Think." *The Atlantic Monthly* 176(1):101–108. http://www.theatlantic.com/magazine/archive/1945/7/as-we-may-think/3881/

Brown, Jesse. 2011. "Stuck between invention and implementation." *Maclean's*, February 25. http://www2.macleans.ca/2011/02/25/stuck-between-invention-and-implementation/

Cameron, Andy. 2009. "Dinner with Myron Or: Rereading Artificial Rality 2: Reflections on Interface and Art." In *ARt&D : research and development in art,* edited by Joke Brouwer, Sandra Fauconnier, Arjen Mulder, Anne Nigten, 10–26. Rotterdam: NAI Publishers.

Cruz-Neira,C., Sandin, D.J., DeFanti, T.A., Kenyon, R.V., and Hart, J.C. 1992. “The CAVE: Audio Visual Experience Automatic Virtual Environment.” *Communications of the ACM*, Vol. 35, No. 6, 65-72.

Engelbart, Douglas. 1962. “Augmenting The Human Intellect: A Conceptual Framework.” *SRI Summary Report AFOSR-3223.* Report prepared for *Director of Information Sciences, Air Force Office of Scientific Research, Washington 25, DC, Contract AF 49(638)–1024 • SRI Project No. 3578 (AUGMENT,3906,).* Online version of the report and original scan can be found here:
(http://www.dougengelbart.org/pubs/augment-3906.html)

Erickson, Thomas D. 1990. “Working With Interface Metaphors.” In *The Art of human-computer interface design*, edited by Brenda Laurel, 65–73. Reading, Mass.: Addison-Wesley.

van Erp, Jan B. F., Ki-Uk Kyung, Sebastian Kassner, Jim Carter, Stephen Brewster, Gerhard Weber and Ian Andrew. 2010. “Setting the Standards for Haptic and Tactile Interactions: ISO’s Work.” *Lecture Notes in Computer Science*, Vol. 6192, 353–358, DOI: 10.1007/978-3-642-14075-4_52

Freeman, Eric Thomas. 1997. “The Lifestreams Software Architecture” PhD Diss., Yale University. http://www.cs.yale.edu/homes/freeman/dissertation/etf.pdf

Gallagher, Shaun. 1995 “Body Schema and Intentionality.” In *The Body and the Self*, edited by J. Bermudez et al., 225–244. Cambridge, MA: MIT Press.

Haraway, Donna. 1991. “A Cyborg Manifesto.” In *Simians, Cyborgs, and Women. The Reinvention of Nature*, edited by D. Haraway, 149–181. London: Free Association Books. (available online: http://www.stanford.edu/dept/HPS/Haraway/CyborgManifesto.html)

Helbig, Hannah B and Marc O. Ernst. 2008. “Haptic Perception in Interaction With Other Senses.” In *Human Haptic Perception: Basics and Applications,* edited by Martin Grunwald, 235–249. Basel: Birkhäuser Basel.

Hsu, Hui-mei Justina. 2011.“The Potential of Kinect as Interactive Educational Technology.” Paper presented at 2nd International Conference on Education and Management Technology, Shanghai, China, August 19-21.

Hutchins, Edwin L., James D. Hollan, and Donald A. Norman. 1985. “Direct Manipulation Interfaces.” *Human-Computer Interaction*, Vol 1, 311–338

Jonas, Hans. 1954. “The Nobility of Sight.” *Philosophy and Phenomenological Research*, Vol. 14, No. 4 , pp. 507–519, http://www.jstor.org/stable/2103230

Kay, Alan. 1990. “User Interface: A Personal View.” In *The Art of human-computer interface design*, edited by Brenda Laurel, 191–207. Reading, Mass.: Addison-Wesley.

Kittler, Friedrich Adolf. 1992. "There is no Software.” *Stanford Literature Review.* 9,1:81–90. (English). http://www.ctheory.net/articles.aspx?id=74

Ishii, Hiroshi and Brygg Ullmer. 1997. “Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms.” Published In Proceedings of CHI '97, March 22-27.

Licklider, J.C.R. 1960. “Man-machine Symbiosis.” *IRE Transactions on Human Factors in Electronics*,
Vol. HFE-1, 4–11. http://groups.csail.mit.edu/medg/people/psz/Licklider.html

MacLean, K. E. 2008. “Haptic Interaction Design for Everyday Interfaces.” *Reviews of Human Factors and Ergonomics* Vol 4, no 1:149–194. doi:10.1518/155723408X342826.

Mann, Steve. 2000. “Telepointer: Hands-Free Completely Self Contained Wearable Visual Augmented Reality without Headwear and without any Infrastructural Reliance,” IEEE International Symposium on Wearable Computing (ISWC00), 177–178. DOI: 10.1109/ISWC.2000.888489

Merrill, D, J. Kalanithi and P. Maes. 2007. “Siftables: Towards Sensor Network User Interfaces.” Presented at First International Conference on Tangible and Embedded Interaction (TEI'07), Baton Rouge, Louisiana, USA, February 15-17.

Nelson, Theodor. 2003. “No More Teacher's Dirty Looks.” In *The New Media Reader*, edited by
Noah Wardrip-Fruin and Nick Montfort, 308–338. Cambridge, Mass. : MIT Press. Originally published in Computer Decisions (September issue, 1970).

Norman, Donald. A and Jacob Nielsen. 2010. “Gestural Interfaces: A step backward in usability.” *Interactions*, Vol. 17, Issue 5, September–October, 46–49. DOI: 10.1145/1836216.1836228

Norman, Donald. 1990. “Why Interfaces Don’t Work.” In *The Art of human-computer interface design*, edited by Brenda Laurel, 209–219. Reading, Mass.: Addison-Wesley.

Penny, Simon. 1994. “Embodied Cultural Agents: at the intersection of Robotics, Cognitive Science and Interactive Art.” AAAI Technical Report FS–97–02, 103–105 http://www.aaai.org/Papers/Symposia/Fall/1997/FS-97-02/FS97-02-024.pdf

de los Reyes, August. 2008. “Predicting the Past.”Presented at Web Directions South 2008. Sydney Convention Centre, Australia, September 25. http://www.webdirections.org/resources/august-de-los-reyes-predicting-the-past/

Robles-De-La-Torres, Gabriel. 2006. “The Importance of the Sense of Touch in Virtual and Real Environments.” *IEEE Multimedia, Special issue on Haptic User Interfaces for Multimedia Systems*, Vol 13, no. 3, 24–30.

Steinberg, Steve G. 1997. “Lifestreams.” *Wired,* February

Turner, Jeremy. 2002. “Myron Krueger Live” *Ctheory*, January 23, http://www.ctheory.net/articles.aspx?id=328

Woods, D.D, E. M. Roth. 1988. “Cognitive Engineering: Human Problem Solving with Tools.” *Human Factors*, Vol. 30(4), 415-430

## VIDEO

Fletcher, Joseph. 2009. “Untold Stories of Touch, Gesture, & NUI or Touch and Gesture Computing, What You Haven't Heard.” Presented at MIX09, Las Vegas, USA, March 21. http://channel9.msdn.com/Events/MIX/MIX09/C15F

Nelson, Ted. 2001. “Zig Zag (Technical briefing).” Presented at ACM Hypertext, University of Aarhus. Århus, Denmark, August 14-18

Barraza Rick, Joshua Blake, Neil Roodyn, Bart Roozendaal, Josh Santangelo, Nicolas Calvi, Dennis Vroegop. 2011. “Microsoft Surface MVPs Present: Natural User Interfaces, Today and Tomorrow; An Interactive Discussion and Demonstration.” Presented at MIX11, Las Vegas, USA, April 12. http://channel9.msdn.com/events/MIX/MIX11/OPN09

## BLOG ENTRIES

Monson-Haefel, Richard. 2008. “Engelbart's Usability Dilemma: Efficiency vs Ease-of-Use.” *Ajax World Magazine,* April 10. http://ajax.sys-con.com/node/536976

Smalley, Eric. 2011. “Kinect makes your hand a quadrocopter remote.” CNet News, July 5. http://news.cnet.com/8301-17938_105-20077017-1/kinect-makes-your-hand-a-quadrocopter-remote/

**WEBPAGES:**

BlablaLab. 2011. “Be Your Own Souvenir.” Accessed November 14. http://byos.blablablab.org/

Electronic Vizualization Library. 2011. “Dream Grrrls.” Accessed October 5. http://www.evl.uic.edu/dolinsky/DG/

Electronic Visualization Laboratory. 2009. Accessed October 16 2011 http://www.evl.uic.edu/core.php?mod=4&type=1&indi=421

Golan Levin. 2011. “Interactive Art & Computational Design / Spring 2011 ”. Accessed August 3. http://golancourses.net/2011spring/projects/project-3-interaction/

Julian Oliver. 2011. “Levelhead.” Accessed September 17. http://julianoliver.com/levelhead

London Science Museum. 2011. “Babbage.” Accessed January 10. http://www.sciencemuseum.org.uk/onlinestuff/stories/babbage.aspx?

Microsoft. 2011. “The Power of PixelsenseTM” Accessed November 15. http://www.microsoft.com/surface/en/us/pixelsense.aspx

Reactable. 2011. “The Reactable.” Accessed November 5. http://www.reactable.com/products/reactable_experience/reactable/

Robust Robotics Group. 2011. “Visual Odometry For GPS-Denied Flight And Mapping Using A Kinect.” Accessed August 31. http://groups.csail.mit.edu/rrg/index.php?n=Main.VisualOdometryForGPS-DeniedFlight

Selling To Schools. 2011. "National Survey of Interactive Whiteboard Usage 2011.” Accessed November 14. https://www.sellingtoschools.com/products/national-survey-interactive-whiteboard-usage-2011

Sixth Sense on Google Code. 2011. “Sixthsense - An open source project that allows people to create their own SixthSense Device and augment the current codebase with their own apps.” Accessed http://code.google.com/p/sixthsense/

Stelarc. 2011. “Ear On Arm.” Accessed October 23.
http://stelarc.org/?catID=20242

Karon MacLean. 2011. Research statement: No title. Accessed July 15.
http://www.cs.ubc.ca/~maclean/