44.
45. Geor...
46. Philip Joh...
Consciousness (Cambr...



2

# The Interface

In 1984 the director of *Blade Runner,* Ridley Scott, was hired to create a commercial to introduce Apple Computer's new Macintosh. In retrospect, this event is full of historical significance. As Peter Lunenfeld has pointed out, *Blade Runner* (1982) and the Macintosh computer (1984)—released within two years of each other—defined the two aesthetics that, twenty years later, still rule contemporary culture, miring us in what he calls the "permanent present." One was a futuristic dystopia which combined futurism and decay, computer technology and fetishism, retro-styling and urbanism, Los Angeles and Tokyo. Since *Blade Runner*'s release, its techno-noir has been replayed in countless films, computer games, novels, and other cultural objects. And although a number of strong aesthetic systems have been articulated in the following decades, both by individual artists (Matthew Barney, Mariko Mori) and by commercial culture at large (the 1980s "postmodern" pastiche, the 1990s techno-minimalism), none of them has been able to challenge the hold of *Blade Runner* on our vision of the future.

In contrast to the dark, decayed, "postmodern" vision of *Blade Runner,* the Graphical User Interface (GUI), popularized by Macintosh, remained true to the modernist values of clarity and functionality. The user's screen was ruled by straight lines and rectangular windows that contained smaller rectangles of individual files arranged in a grid. The computer communicated with the user via rectangular boxes containing clean black type rendered against a white background. Subsequent versions of GUI added colors and made it possible for users to customize the appearance of many interface elements, thus somewhat diluting the sterility and boldness of the original monochrome 1984 version. Yet its original aesthetic survives in the displays of hand-held communicators such as Palm Pilot, cellular telephones, car navigation systems, and other consumer electronic products that use small LCD displays comparable in quality to the 1984 Macintosh screen.

Like *Blade Runner,* Macintosh's GUI articulated a vision of the future, although a very different one. In this vision, the lines between the human and its technological creations (computers, androids) are clearly drawn, and decay is not tolerated. In a computer, once a file is created, it never disappears except when explicitly deleted by the user. And even then deleted items can usually be recovered. Thus, if in "meatspace" we have to work to remember, in cyberspace we have to work to forget. (Of course while they run, OS and applications constantly create, write to, and erase various temporary files, as well as swap data between RAM and virtual memory files on a hard drive, but most of this activity remains invisible to the user.)

Also like *Blade Runner,* GUI vision came to influence many other areas of culture. This influence ranges from the purely graphical (for instance, the use of GUI elements by print and TV designers) to the more conceptual. In the 1990s, as the Internet progressively grew in popularity, the role of the digital computer shifted from being a particular technology (a calculator, symbol processor, image manipulator, etc.) to a filter for all culture, a form through which all kinds of cultural and artistic production were mediated. As the window of a Web browser replaced cinema and television screen, the art gallery wall, library and book, all at once, the new situation manifested itself: All culture, past and present, came to be filtered through a computer, with its particular human-computer interface.[1]

In semiotic terms, the computer interface acts as a code that carries cultural messages in a variety of media. When you use the Internet, everything you access—texts, music, video, navigable spaces—passes through the interface of the browser and then, in turn, the interface of the OS. In cultural communication, a code is rarely simply a neutral transport mechanism; usually it affects the messages transmitted with its help. For instance, it may make some messages easy to conceive and render others unthinkable. A code may also provide its own model of the world, its own logical system, or ideology; subsequent cultural messages or whole languages created with this code will be limited by its accompanying model, system, or ideology. Most modern cultural theories rely on these notions, which together I will refer to as the "non-transparency of the code" idea. For instance, according to the Whorf-Sapir hypothesis, which enjoyed popularity in the middle of the twentieth century, human thinking is determined by the code of natural language; the speakers of different natural languages perceive and think about the world differently.[2] The Whorf-Sapir hypothesis is an extreme expression of the "non-transparency of the code" idea; usually it is formulated in less extreme forms. But when we think about the case of the human-computer interface, applying a "strong" version of this idea makes sense. The interface

1. Stephen Johnson's *Interface Culture* makes a claim for the cultural significance of computer interface.

2. Other examples of cultural theories that rely on the "non-transparency of the code" idea are Yuri Lotman's theory of secondary modeling systems, George Lakoff's cognitive linguistics, Jacques Derrida's critique of logocentrism, and Marshall McLuhan's media theory.



shapes how the computer user conceives of the computer itself. It also determines how users think of any media object accessed via a computer. Stripping different media of their original distinctions, the interface imposes its own logic on them. Finally, by organizing computer data in particular ways, the interface provides distinct models of the world. For instance, a hierarchical file system assumes that the world can be organized in a logical multilevel hierarchy. In contrast, a hypertext model of the World Wide Web arranges the world as a nonhierarchical system ruled by metonymy. In short, far from being a transparent window into the data inside a computer, the interface brings with it strong messages of its own.

As an example of how the interface imposes its own logic on media, consider "cut and paste" operations, standard in all software running under the modern GUI. This operation renders insignificant the traditional distinction between spatial and temporal media, since the user can cut and paste parts of images, regions of space, and parts of a temporal composition in exactly the same way. It is also "blind" to traditional distinctions in scale: the user can cut and paste a single pixel, an image, or a whole digital movie in the same way. And last, this operation also renders insignificant the traditional distinctions between media: "cut and paste" can be applied to texts, still and moving images, sounds, and 3-D objects in the same way.

The interface comes to play a crucial role in the information society in yet another way. In this society, work and leisure activities not only increasingly involve computer use, but they also converge around the same interfaces. Both "work" applications (word processors, spreadsheet programs, database programs) and "leisure" applications (computer games, informational DVD) use the same tools and metaphors of GUI. The best example of this convergence is a Web browser employed both in the office and at home, both for work and for play. In this respect information society is quite different from industrial society, with its clear separation between the field of work and the field of leisure. In the nineteenth century Karl Marx imagined that a future communist state would overcome this work-leisure divide as well as the highly specialized and piecemeal character of modern work itself. Marx's ideal citizen would be cutting wood in the morning, gardening in the afternoon, and composing music in the evening. Today, the subject of the information society is engaged in even more activities during a typical day: inputting and analyzing data, running simulations, searching the Internet, playing computer games, watching streaming video, listening to music online, trading stocks, and so on. Yet in performing all these different activities, the user in essence is always using the same few tools and commands: a computer screen and a mouse; a Web browser; a search engine; cut, paste, copy, delete, and find commands.

If the human-computer interface has become a key semiotic code of the information society as well as its metatool, how does this affect the functioning of cultural objects in general and art objects in particular? As I have already noted, in computer culture it becomes common to construct a number of different interfaces to the same "content." For instance, the same data can be represented as a 2-D graph or as an interactive navigable space. Or, a Web site may guide the user to different versions of the site depending on the bandwidth of her Internet connection. Given these examples, we may be tempted to think of a new media artwork as also possessing two separate levels: content and interface. Thus, the old dichotomies *content—form* and *content—medium* can be rewritten as *content—interface*. But postulating such an opposition assumes that that artwork's content is independent of its medium (in an art historical sense) or its code (in a semiotic sense). Situated in some idealized medium-free realm, content is assumed to exist before its material expression. These assumptions are correct in the case of the visualization of quantified data; they also apply to classical art with its well-defined iconographic motives and representational conventions. But just as modern thinkers, from Whorf to Derrida, insisted on the "nontransparency of the code" idea, modern artists assumed that content and form cannot be separated. In fact, from the "abstraction" of the 1910s to the "process" of the 1960s, artists have continued to invent concepts and procedures to assure the impossibility of painting some preexistent content.

This leaves us with an interesting paradox. Many new media artworks have what can be called an "informational dimension," the condition that they share with all new media objects. The experience includes retrieving, looking at and thinking about quantified data. Therefore, when we refer to such artworks, we are justified in separating the levels of content and interface. At the same time, new media artworks have more traditional "experiential" or aesthetic dimensions, which justify their status as art rather than information design. These dimensions include a particular configuration of space, time, and surface articulated in the work; a particular sequence of the user's activities over time in interacting with the work; a particular formal, material, and phenomenological user experience. And it is the work's interface that creates its unique materiality and a unique user experience. To change the interface even slightly is to change the work dramatically. From this perspective, to think of an interface as a separate level, as something that can be arbitrarily varied, is to eliminate the status of a new media artwork as art.



There is another way to think about the difference between new media design and new media art in relation to the content—interface dichotomy. In contrast to design, in art the connection between content and form (or, in the case of new media, content and interface) is motivated; that is, the choice of a particular interface is motivated by a work's content to such degree that it can no longer be thought of as a separate level. Content and interface merge into one entity, and no longer can be taken apart.

Finally, the idea of content preexisting interface is challenged in yet another way by new media artworks that dynamically generate their data in real time. While in a menu-based interactive multimedia application or a static Web site, all data already exists before the user accesses it, in dynamic new media artworks, the data is created on the fly, or, to use the new media lingo, at run time. This can be accomplished in a variety of ways: procedural computer graphics, formal language systems, AI and AL programming. All these methods share the same principle: a programmer sets up some initial conditions, rules, or procedures that control the computer program generating the data. For the purposes of the present discussion, the most interesting of these approaches are AL and the evolution paradigm. In the AL approach, the interaction between a number of simple objects at run time leads to the emergence of complex global behaviors. These behaviors can only be obtained in the course of running the computer program; they cannot be predicted beforehand. The evolution paradigm applies the metaphor of evolution theory to the generation of images, shapes, animations, and other media data. The initial data supplied by the programmer acts as a genotype that is expanded into a full phenotype by the computer. In either case, the content of an artwork is the result of a collaboration between the artist/programmer and the computer program, or, if the work is interactive, between the artist, the computer program, and the user. New media artists who have most systematically explored the AL approach are the team of Christa Sommerer and Laurent Mignonneau. In their installation "Life Spacies," virtual organisms appear and evolve in response to the position, movement, and interactions of visitors. Artist/programmer Karl Sims also made key contributions to applying the evolution paradigm to media generation. In his installation "Galapagos" computer programs generate twelve different virtual organisms at every iteration; visitors select an organism that will continue to live, copulate, mutate, and reproduce.[3] Commercial products that use AL and evolution approaches include computer games such as the *Creatures* series (Mindscape Entertainment) and "virtual pet" toys such as Tamagochi.

In organizing this book, I wanted to highlight the importance of the interface category by placing its discussion right in the beginning. The two sections of this chapter present examples of different issues raised by this category—but they in no way exhaust it. In "The Language of Cultural Interface," I introduce the term "cultural interfaces" to describe interfaces used by stand-alone hypermedia (CD-ROM and DVD titles), Web sites, computer games, and other cultural objects distributed via computers. I analyze how the three cultural forms of the cinema, the printed word, and a general-purpose human-computer interface contributed to shaping the appearance and functionality of cultural interfaces during the 1990s.

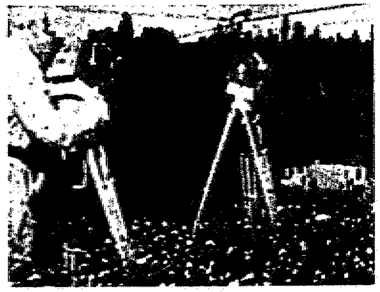
The second section, "The Screen and the User," discusses the key element of the modern interface—the computer screen. As in the first section, I am interested in analyzing continuities between the computer interface and older cultural forms, languages, and conventions. This section positions the computer screen within a larger historical tradition and traces different stages in the development of this tradition—the static illusionistic image of Renaissance painting; the moving image of the film screen; the real-time image of radar and television; and the real-time interactive image of the computer screen.

---

3. http://www.ntticc.or.jp/permanent/index_e.html.



## The Language of Cultural Interfaces



### Cultural Interfaces

The term *human-computer interface* describes the ways in which the user interacts with a computer. HCI includes physical input and output devices such as a monitor, keyboard, and mouse. It also consists of metaphors used to conceptualize the organization of computer data. For instance, the Macintosh interface introduced by Apple in 1984 uses the metaphor of files and folders arranged on a desktop. Finally, HCI also includes ways of manipulating data, that is, a grammar of meaningful actions that the user can perform on it. Examples of actions provided by modern HCI are copy, rename, and delete a file; list the contents of a directory; start and stop a program; set the computer's date and time.

The term HCI was coined when the computer was used primarily as a tool for work. However, during the 1990s, the identity of the computer changed. In the beginning of the decade, the computer was still largely thought of as a simulation of a typewriter, paintbrush or drafting ruler—in other words, as a tool used to produce cultural content that, once created, would be stored and distributed in the appropriate media—printed page, film, photographic print, electronic recording. By the end of the decade, as Internet use became commonplace, the computer's public image was no longer solely that of a tool but also a universal media machine, which could be used not only to author, but also to store, distribute, and access all media.

As distribution of all forms of culture becomes computer-based, we are increasingly "interfacing" to predominantly cultural data—texts, photographs, films, music, virtual environments. In short, we are no longer interfacing to a computer but to culture encoded in digital form. I will use the term *cultural interface* to describe a human-computer-culture interface—the ways in which computers present and allow us to interact with cultural data. Cultural interfaces include the interfaces used by the designers of Web sites, CD-ROM and DVD titles, multimedia encyclopedias, on-line museums and magazines, computer games, and other new media cultural objects.

If you need to remind yourself what a typical cultural interface looked like in the second part of the 1990s, say 1997, go back in time and click to a random Web page. You are likely to see something that graphically resembles a magazine layout from the same decade. The page is dominated by text—headlines, hyperlinks, blocks of copy. Within this text are a few media elements—graphics, photographs, perhaps a QuickTime movie, and a VRML scene. The page also includes radio buttons and a pull-down menu that allows you to choose an item from the list. Finally, there is a search engine: Type a word or a phrase, hit the search button, and the computer will scan through a file or database trying to match your entry.

For another example of a prototypical cultural interface of the 1990s, you might load (assuming it would still run on your computer) the most well-known CD-ROM of the 1990s—*Myst* (Broderbund, 1993). Its opening clearly recalls a movie: credits slowly scroll across the screen, accompanied by a movie-like soundtrack to set the mood. Next, the computer screen shows an open book, awaiting the click of a mouse. Next, a familiar element of a Macintosh interface makes an appearance, reminding you that besides being a new movie/book hybrid, *Myst* is also a computer application: you can adjust the sound volume and graphics quality by selecting from a standard Macintosh-style menu at the upper top of the screen. Finally, you are taken inside the game, where the interplay between the printed word and cinema continues. A virtual camera frames images of an island that dissolve between each other. At the same time, you keep encountering books and letters, which take over the screen, providing with you with clues on how to progress in the game.

Given that computer media is simply a set of characters and numbers stored in a computer, there are numerous ways in which it could be presented to a user. Yet, as is the case with all cultural languages, only a few of these possibilities actually appear viable at any given historical moment. Just as early fifteenth-century Italian painters could only conceive of painting in a very particular way—quite different from, say, sixteenth-century Dutch



painters—today's digital designers and artists use only a small set of action grammars and metaphors out of a much larger set of all possibilities.

Why do cultural interfaces—Web pages, CD-ROM titles, computer games—look the way they do? Why do designers organize computer data in certain ways and not in others? Why do they employ some interface metaphors and not others?

In my view, the language of cultural interfaces is largely made up from elements of other, already familiar cultural forms. In the following I will explore the contributions of three such forms to this language during its first decades—the 1990s. The three forms on which I will focus make their appearance in the opening sequence of the already discussed prototypical new media object of the 1990s—*Myst*. Its opening activates them before our eyes, one by one. The first form is cinema. The second is the printed word. The third is a general-purpose human-computer interface.

As should become clear, I use "cinema" and "printed word" as shortcuts. They stand not for particular objects, such as a film or a novel, but rather for larger cultural traditions (we can also use such terms as "cultural forms," "mechanisms," "languages," or "media"). "Cinema" thus includes the mobile camera, representations of space, editing techniques, narrative conventions, spectator activity—in short, different elements of cinematic perception, language, and reception. Their presence is not limited to the twentieth-century institution of fiction films; they can be found already in panoramas, magic lantern slides, theater, and other nineteenth-century cultural forms; similarly, since the middle of the twentieth century, they have been present not only in films but also in television and video programs. In the case of the "printed word," I am also referring to a set of conventions that have developed over many centuries (some even before the invention of print) and that today are shared by numerous forms of printed matter, from magazines to instruction manuals—a rectangular page containing one or more columns of text, illustrations or other graphics framed by the text, pages that follow each other sequentially, a table of contents, and index.

The modern human-computer interface has a much shorter history than the printed word or cinema—but it is still a history. Principles such as direct manipulation of objects on the screen, overlapping windows, iconic representation, and dynamic menus were gradually developed over a few decades, from the early 1950s to the early 1980s, when they finally appeared in commercial systems such as Xerox Star (1981), the Apple Lisa (1982), and most importantly the Apple Macintosh (1984).[4] Since then, they have become accepted conventions for operating a computer, and a cultural language in its own right.

Cinema, the printed word, the human-computer interface: Each of these traditions has developed its own unique way of organizing information, presenting it to the user, correlating space and time, and structuring human experience in the process of accessing information. Pages of text and a table of contents; 3-D spaces framed by a rectangular frame that can be navigated using a mobile point of view; hierarchical menus, variables, parameters, copy/paste and search/replace operations—these and other elements of the three traditions are shaping cultural interfaces today. Cinema, the printed word, and HCI are the three main reservoirs of metaphors and strategies for organizing information which feed cultural interfaces.

Treating them as if they occupied the same conceptual plane has an advantage—a theoretical bonus. It is only natural to think of them as belonging to two different kinds of cultural species, so to speak. If HCI is a general purpose tool which can be used to manipulate any kind of data, both the printed word and cinema are less general, and offer their own ways to organize particular types of data: text in the case of print, audio-visual narrative taking place in a 3-D space in the case of cinema. HCI is a system of controls to operate a machine; the printed word and cinema are cultural traditions, distinct ways of recording human memory and human experience, mechanisms for the cultural and social exchange of information. Bringing HCI, the printed word, and cinema together allows us to see that the three have more in common than we might have anticipated. On the one hand, being part of our culture now for half a century, HCI already represents a powerful cultural tradition, a cultural language offering its own ways of representing human memory and human experience. This language speaks in the form of discrete objects organized in hierarchies (hierarchical file system), or as catalogs (databases), or as objects linked together through hyperlinks (hypermedia). On the other hand, we begin to see that the printed word and cinema

4. Brad A. Myers, "A Brief History of Human Computer Interaction Technology," technical report CMU-CS-96-163 and Human Computer Interaction Institute Technical Report CMU-HCII-96-103 (Pittsburgh, Pa.: Carnegie Mellon University, Human-Computer Interaction Institute, 1996).



also can be thought of as interfaces, even though historically they have been tied to particular kinds of data. Each has its own grammar of actions, each comes with its own metaphors, each offers a particular physical interface. A book or a magazine is a solid object consisting of separate pages; actions include going from page to page linearly, marking individual pages, and using the table of contents. In the case of cinema, its physical interface is the particular architectural arrangement of the movie theater; its metaphor, a window opening up into a virtual 3-D space.

Today, as media is being "liberated" from traditional physical storage media—paper, film, stone, glass, magnetic tape—elements of the printed word interface and the cinema interface that previously were hardwired to content become "liberated" as well. A digital designer can freely mix pages and virtual cameras, tables of content and screens, bookmarks and points of view. No longer embedded within particular texts and films, these organizational strategies are now free floating in our culture, available for use in new contexts. In this respect, the printed word and cinema have indeed become interfaces—rich sets of metaphors, ways of navigating through content, ways of accessing and storing data. For a computer user, both conceptually and psychologically, their elements exist on the same plane as radio buttons, pull-down menus, command line calls, and other elements of the standard human-computer interface.

Let us now discuss some of the elements of these three cultural traditions—cinema, the printed word, and HCI—to see how they have shaped the language of cultural interfaces.

### Printed Word

In the 1980s, as PCs and word processing software became commonplace, text became the first cultural medium to be subjected to digitization in a massive way. Already in the 1960s, two and a half decades before the concept of digital media was born, researchers were thinking about making the sum total of human written production—books, encyclopedias, technical articles, works of fiction, and so on—available online (Ted Nelson's *Xanadu* project[5]).

5. http://www.xanadu.net.

Text is unique among media types. It plays a privileged role in computer culture. On the one hand, it is one media type among others. But, on the other hand, it is a metalanguage of computer media, a code in which all other media are represented: coordinates of 3-D objects, pixel values of digital images, the formatting of a page in HTML. It is also the primary means of communication between a computer and a user: One types single line commands or runs computer programs written in a subset of English; the other responds by displaying error codes or text messages.[6]

If computers use text as their metalanguage, cultural interfaces in their turn inherit the principles of text organization developed by human civilization throughout its existence. One of these principles is a page—a rectangular surface containing a limited amount of information, designed to be accessed in some order, and having a particular relationship to other pages. In its modern form, the page was born in the first centuries of the Christian era when the clay tablet and papyrus roll were replaced by the codex—a collection of written pages stitched together on one side.

Cultural interfaces rely on our familiarity with the "page interface" while also trying to stretch its definition to include new concepts made possible by the computer. In 1984, Apple introduced a graphical user interface that presented information in overlapping windows stacked behind one another—essentially, a set of book pages. The user was given the ability to go back and forth between pages, as well as to scroll through individual pages. In this way, a traditional page was redefined as a virtual page, a surface that can be much larger than the limited surface of a computer screen. In 1987, Apple introduced the popular *Hypercard* program, which extended the page concept in new ways. Now, users were able to include multimedia elements within pages, as well as to establish links between pages regardless of their ordering. A few years later, designers of HTML stretched the concept of a page even further by enabling the creation of distributed documents; that is, different parts of a document are located on different computers connected through the network. With this development, a long process of gradual "virtualization" of the page reached a new stage. Messages written on clay tablets, which were almost indestructible, were replaced by ink on paper. Ink, in its turn, was replaced by bits of computer memory, making characters on an electronic screen. Now, with HTML, which allows parts of a single page to be located on different computers, the page becomes even more fluid and unstable.

6. XML, which is promoted as the replacement for HTML, enables any user to create her own customized markup language. The next stage in computer culture may involve authoring not simply new Web documents but new languages. For more information on XML, see http://www.ucc.ie/xml.



The conceptual development of the page in computer media can also be read in a different way—not as a further development of a codex form, but as a return to earlier forms such as the papyrus roll of ancient Egypt, Greece, and Rome. Scrolling through the contents of a computer window or a World Wide Web page has more in common with unrolling than it does with turning the pages of a modern book. In the case of the Web of the 1990s, the similarity with a roll is even stronger because information is not available all at once, but rather arrives sequentially, top to bottom.

A good example of how cultural interfaces stretch the definition of a page while mixing together its different historical forms is the Web page created in 1997 by the British design collective antirom for HotWired's RGB Gallery.[7] The designers created a large surface containing rectangular blocks of text in different font sizes, arranged without any apparent order. The user is invited to skip from one block to another moving in any direction. Here, the different directions of reading used in different cultures are combined together on a single page.

By the mid-1990s, Web pages included a variety of media types—but they were still essentially traditional pages. Different media elements—graphics, photographs, digital video, sound, and 3-D worlds—were embedded within rectangular surfaces containing text. To this extent, a typical Web page was conceptually similar to a newspaper page, which is also dominated by text, with photographs, drawings, tables, and graphs embedded in between, along with links to other pages of the newspaper. VRML evangelists wanted to overturn this hierarchy by imaging in a future in which the World Wide Web is rendered as a giant 3-D space, with all the other media types, including text, existing within it.[8] Given that the history of a page

7. http://www.hotwired.com/rgb/antirom/index2.html.

8. See, for instance, Mark Pesce, "Ontos, Eros, Noos, Logos," the keynote address for the International Symposium on Electronic Arts (ISEA), 1995, http://www.xs4all.nl/~mpesce/iseakey.html.

stretches back for thousands of years, I think it is unlikely that it will disappear so quickly.

As the Web page became a new cultural convention, its dominance was challenged by two Web browsers created by artists—Web Stalker (1997) by the I/O/D collective[9] and Netomat (1999) by Maciej Wisniewski.[10] Web Stalker emphasizes the hypertextual nature of the Web. Instead of rendering standard Web pages, it renders the networks of hyperlinks these pages embody. When a user enters a URL for a particular page, Web Stalker displays all pages linked to that page as a line graph. Netomat similarly refuses the page convention of the Web. The user enters a word or a phrase that is passed to search engines. Netomat then extracts page titles, images, audio, or any other media type, as specified by the user, from the found pages and floats them across the computer screen. As can be seen, both browsers refuse the page metaphor, instead substituting their own metaphors—a graph showing the structure of links in the case of Web Stalker, a flow of media elements in the case of Netomat.

While the 1990s' Web browsers and other commercial cultural interfaces have retained the modern page format, they also have come to rely on a new way of organizing and accessing texts that has little precedent within the book tradition—hyperlinking. We may be tempted to trace hyperlinking to earlier forms and practices of non-sequential text organization, such as the Torah's interpretations and footnotes, but it is actually fundamentally different from them. Both the Torah's interpretations and footnotes imply a master-slave relationship between one text and another. But in the case of hyperlinking as implemented by HTML and earlier by Hypercard, no such relationship of hierarchy is assumed. The two sources connected through a hyperlink have equal weight; neither one dominates the other. Thus the acceptance of hyperlinking in the 1980s can be correlated with contemporary culture's suspicion of all hierarchies, and preference for the aesthetics of collage in which radically different sources are brought together within a singular cultural object.

Traditionally, texts encoded human knowledge and memory, instructed, inspired, convinced, and seduced their readers to adopt new

9. http://www.backspace.org/iod.

10. http://www.netomat.net.



ideas, new ways of interpreting the world, new ideologies. In short, the printed word was linked to the art of rhetoric. While it is probably possible to invent a new rhetoric of hypermedia that will use hyperlinking not to distract the reader from the argument (as is often the case today), but rather to further convince her of an argument's validity, the sheer existence and popularity of hyperlinking exemplifies the continuing decline of the field of rhetoric in the modern era. Ancient and medieval scholars classified hundreds of different rhetorical figures. In the middle of the twentieth century, linguist Roman Jakobson, under the influence of the computer's binary logic, information theory, and cybernetics to which he was exposed at MIT where he was teaching, radically reduced rhetoric to just two figures—metaphor and metonymy.[11] Finally, in the 1990s, World Wide Web hyperlinking has privileged the single figure of metonymy at the expense of all others.[12] The hypertext of the World Wide Web leads the reader from one text to another, ad infinitum. Contrary to popular images of computer media as collapsing all human culture into a single giant library (which implies the existence of some ordering system), or a single giant book (which implies a narrative progression), it is perhaps more accurate to think of the new media culture as an infinite flat surface where individual texts are placed in no particular order, like the Web page designed by antirom for HotWired. Expanding this comparison further, we can note that Random Access Memory, the concept behind the group's name, also implies a lack of hierarchy: Any RAM location can be accessed as quickly as any other. In contrast to the older storage media of book, film, and magnetic tape, where data is organized sequentially and linearly, thus suggesting the presence of a narrative or a rhetorical trajectory, RAM "flattens" the data. Rather than seducing the user through a careful arrangement of arguments and examples, points and counterpoints, changing rhythms of presentation (i.e., the rate of data streaming, to use contemporary language), simulated false paths, and dramatically presented conceptual

11. Roman Jakobson, "Deux aspects du langage et deux types d'aphasie," in *Temps Modernes*, no. 188 (January 1962).

12. XLM diversifies types of links available by including bidirectional links, multiway links, and links to a span of text rather than a simple point.

breakthroughs, cultural interfaces, like RAM itself, bombard the user with all the data at once.[13]

In the 1980s many critics described one of the key effects of "postmodernism" as that of spatialization—privileging space over time, flattening historical time, refusing grand narratives. Computer media, which evolved during the same decade, accomplished this spatialization quite literally. It replaced sequential storage with random-access storage; hierarchical organization of information with a flattened hypertext; psychological movement of narrative in novels and cinema with physical movement through space, as witnessed by endless computer animated fly-throughs or computer games such as *Myst*, *Doom*, and countless others. In short, time became a flat image or a landscape, something to look at or navigate through. If there is a new rhetoric or aesthetic possible here, it may have less to do with the ordering of time by a writer or an orator, and more with spatial wandering. The hypertext reader is like Robinson Crusoe, walking across the sand, picking up a navigation journal, a rotten fruit, an instrument whose purpose he does not know; leaving imprints that, like computer hyperlinks, follow from one found object to another.

### Cinema

The printed word tradition that initially dominated the language of cultural interfaces is becoming less important, while the part played by cinematic elements is becoming progressively stronger. This is consistent with a general trend in modern society toward presenting more and more information in the form of time-based audiovisual moving image sequences, rather than as text. As new generations of both computer users and computer designers grow up in a media-rich environment dominated by television rather than by printed texts, it is not surprising that they favor cinematic language over the language of print.

A hundred years after cinema's birth, cinematic ways of seeing the world, of structuring time, of narrating a story, of linking one experience to the

13. This may imply that new digital rhetoric may have less to do with arranging information in a particular order and more to do simply with selecting what is included and what is not included in the total corpus presented.



next, have become the basic means by which computer users access and interact with all cultural data. In this respect, the computer fulfills the promise of cinema as a visual Esperanto—a goal that preoccupied many film artists and critics in the 1920s, from Griffith to Vertov. Indeed, today millions of computer users communicate with each other through the same computer interface. And in contrast to cinema where most "users" are able to understand cinematic language but not speak it (i.e., make films), all computer users can speak the language of the interface. They are active users of the interface, employing it to perform many tasks: send e-mail, organize files, run various applications, and so on.

The original Esperanto never became truly popular. Cultural interfaces, in contrast, are widely used and easily learned. We have what is an unprecedented situation in the history of cultural languages—a language designed by a rather small group of people that is immediately adopted by millions of computer users. How is it possible that people around the world adopt today something that a twenty-something programmer in Northern California hacked together just the night before? Shall we conclude that we are somehow biologically "wired" to the interface language, in the same way as we are "wired" to different natural languages according to the original hypothesis of Noam Chomsky?

The answer is of course no. Users are able to acquire new cultural languages, whether cinema a hundred years ago, or cultural interfaces today, because these languages are based on previous and already familiar cultural forms. In the case of cinema, the cultural forms that went into its making include theater, magic lantern shows, and other nineteenth-century forms of public entertainment. Cultural interfaces in turn draw on older cultural forms such as cinema and the printed word. I have already discussed some ways in which the printed word tradition structures interface language; now it is cinema's turn.

I will begin with probably the most important case of cinema's influence on cultural interfaces—the mobile camera. Originally developed as part of 3-D computer graphics technology for such applications as computer-aided design, flight simulators, and computer movie making, during the 1980s and 1990s the camera model became as much of an interface convention as scrollable windows or cut-and-paste operations. It became an accepted way of interacting with any data represented in three dimensions—which in computer culture means literally anything and everything—the results of a physical simulation, an architectural site, the design of a new molecule, statistical data, the structure of a computer network, and so on. As computer culture gradually spatializes all representations and experiences, they are subjected to the camera's particular grammar of data access. Zoom, tilt, pan, and track—we now use these operations to interact with data spaces, models, objects, and bodies.

Abstracted from its historical temporary "imprisonment" within the physical body of a movie camera directed at physical reality, a virtualized camera also becomes an interface to all types of media and information beside 3-D space. As an example, consider the GUI of the leading computer animation software—PowerAnimator from Alias/Wavefront.[14] In this interface, each window, regardless of whether it displays a 3-D model, a graph, or even plain text, contains Dolly, Track, and Zoom buttons. It is particularly important that the user is expected to dolly and pan over text as if it were a 3-D scene. In this interface, cinematic vision triumphs over the print tradition, with the camera subsuming the page. The Gutenberg galaxy turns out to be just a subset of the Lumières' universe.

Another feature of cinematic perception that persists in cultural interfaces is a rectangular framing of represented reality.[15] Cinema itself inherited this framing from Western painting. Since the Renaissance, the frame has acted as a window onto a larger space that is assumed to extend beyond the frame. This space is cut by the frame's rectangle into two parts: "onscreen space," the part that is inside the frame, and the part that is outside. In the

14. See http://www.aw.sgi.com/pages/home/pages/products/pages/poweranimator_film_sgi/.

15. In *The Address of the Eye,* Vivian Sobchack discusses the three metaphors of frame, window, and mirror that underlie modern film theory. The metaphor of the frame comes from modern painting and is central to formalist theory, which is concerned with signification. The metaphor of the window underlies realist film theory (Bazin), which stresses the act of perception. Realist theory follows Alberti in conceptualizing the cinema screen as a transparent window onto the world. Finally, the metaphor of the mirror is central to psychoanalytic film theory. In terms of these distinctions, my discussion here is concerned with the window metaphor. The distinctions themselves, however, open up a very productive space for thinking further about the relationships between cinema and computer media, in particular, the cinema screen and the computer window. See Vivian Sobchack, *The Address of the Eye: A Phenomenology of Film Experience* (Princeton, N.J.: Princeton University Press, 1992).



famous formulation of Leon Battista Alberti, the frame acts as a window onto the world. Or, in the more recent formulation of French film theorist Jacques Aumont and his co-authors, "The onscreen space is habitually perceived as included within a more vast scenographic space. Even though the onscreen space is the only visible part, this larger scenographic part is nonetheless considered to exist around it."[16]

Just as a rectangular frame in painting and photography presents a part of a larger space outside it, a window in HCI presents a partial view of a larger document. But if in painting (and later in photography), the framing chosen by an artist is final, computer interface benefits from a new invention introduced by cinema—the mobility of the frame. Just as a kino-eye can move around a space revealing its different regions, a computer user can scroll through a window's contents.

It is not surprising to see that screen-based interactive 3-D environments, such as VRML words, also use cinema's rectangular framing, since they rely on other elements of cinematic vision, specifically, a mobile virtual camera. It may be surprising, however, to realize that the Virtual Reality interface, often promoted as the most "natural" interface of all, utilizes the same framing.[17] As in cinema, the world presented to a VR user is cut by a rectangular frame. As in cinema, this frame presents a partial view of a larger space.[18] As in cinema, the virtual camera moves around to reveal different parts of this space.

Of course, the camera is now controlled by the user and in fact is identified with her own sight. Yet it is crucial that in VR one sees the virtual world through a rectangular frame, and that this frame always presents only part of

16. Jacques Aumont et al., *Aesthetics of Film* (Austin: University of Texas Press, 1992), 13.

17. By VR interface, I mean the common forms of a head-mounted or head-coupled directed display employed in VR systems. For a popular review of such displays written when the popularity of VR was at its peak, see Steve Aukstakalnis and David Blatner, *Silicon Mirage: The Art and Science of Virtual Reality* (Berkeley, CA: Peachpit Press, 1992), pp. 80–98. For a more technical treatment, see Dean Kocian and Lee Task, "Visually Coupled Systems Hardware and the Human Interface," in *Virtual Environments and Advanced Interface Design*, ed. Woodrow Barfield and Thomas Furness III (New York and Oxford: Oxford University Press, 1995), 175–257.

18. See Kocian and Task for details on the field of view of various VR displays. Although it varies widely between different systems, the typical size of the field of view in commercial head-mounted displays (HMD) available in the first part of the 1990s was thirty to fifty degrees.

a larger whole. This frame creates a distinct subjective experience that is much closer to cinematic perception than it is to unmediated sight.

Interactive virtual worlds, whether accessed through a screen-based or VR interface, are often discussed as the logical successor to cinema and potentially the key cultural form of the twenty-first century just as cinema was the key cultural form of the twentieth century. These discussions usually focus on issues of interaction and narrative; thus, the typical scenario for twenty-first century cinema involves a user represented as an avatar existing literally "inside" the narrative space, rendered with photorealistic 3-D computer graphics, interacting with virtual characters and perhaps other users, and affecting the course of narrative events.

It is an open question whether this and similar scenarios indeed represent an extension of cinema, or if they rather should be thought of as a continuation of theatrical traditions such as improvisational or avant-garde theater. But what undoubtedly can be observed is how virtual technology's dependence on cinema's mode of seeing and language is becoming progressively stronger. This coincides with the move from proprietary and expensive VR systems to more widely available and standardized technologies, such as VRML. (The following examples refer to a particular VRML browser—WebSpace Navigator 1.1 from SGI.[19] Other VRML browsers have similar features.)

The creator of a VRML world can define a number of viewpoints that are loaded with the world.[20] These viewpoints automatically appear in a special menu in a VRML browser that allows the user to step through them, one by one. Just as in cinema, ontology is coupled with epistemology: the world is designed to be viewed from particular points of view. The designer of a virtual world is thus a cinematographer as well as an architect. The user can wander around the world, or she can save time by assuming the familiar position of a cinema viewer for whom the cinematographer has already chosen the best viewpoints.

Equally interesting is another option that controls how a VRML browser moves from one viewpoint to the next. By default, the virtual camera travels smoothly through space from the current viewpoint to the next as though on a dolly, its movement automatically calculated by the software. Selecting the "jump cuts" option makes it cut from one view to the next. Both modes are obviously derived from cinema. Both are more efficient than trying to explore the world on its own.

With a VRML interface, nature is firmly subsumed under culture. The eye is subordinated to the kino-eye. The body is subordinated to the virtual body of the virtual camera. While the user can investigate the world on her own, freely selecting trajectories and viewpoints, the interface privileges cinematic perception—cuts, precomputed, dolly-like motions, preselected viewpoints.

The area of computer culture where the cinematic interface is being transformed into a cultural interface most aggressively is computer games. By the 1990s, game designers had moved from two to three dimensions and had begun to incorporate cinematic language in an increasingly systematic fashion. Games began to feature lavish opening cinematic sequences (called "cinematics" in the game business) that set the mood, established the setting, and introduced the narrative. Frequently, the whole game would be structured as an oscillation between interactive fragments requiring the user's input and noninteractive cinematic sequences, that is, "cinematics." As the decade progressed, game designers created increasingly complex—and increasingly cinematic—interactive virtual worlds. Regardless of a game's genre, it came to rely on cinematography techniques borrowed from traditional cinema, including the expressive use of camera angles and depth of field, and dramatic lighting of 3-D computer-generated sets to create mood and atmosphere. In the beginning of the decade, many games such as *The 7th Guest* (Trilobyte, 1993) or *Voyeur* (Philips Interactive Media, 1994) used digital video of actors superimposed over 2-D or 3-D backgrounds; by its end, they had switched to fully synthetic characters rendered in real time.[21] This switch allowed game designers to go beyond the branching-type structure of earlier games based on digital video in which all possible scenes had to be taped beforehand. In contrast, 3-D characters animated in real time move

19. http://webspace.sgi.com/WebSpace/Help/1.1/.

20. See John Hartman and Josie Wernecke, *The VRML 2.0 Handbook: Building Moving Worlds on the Web* (Reading, Mass.: Addison-Wesley, 1996), 363.

21. Examples of the earlier trend are *Return to Zork* (Activision, 1993) and *The 7th Guest* (Trilobyte/Virgin Games, 1993). Examples of the later trend are *Soulblade* (Namco, 1997) and *Tomb Raider* (Eidos, 1996).

arbitrarily around the space, and the space itself can change during the game. (For instance, when a player returns to an already visited area, she will find any objects that she left there earlier.) This switch also made virtual words more cinematic, as characters could be better visually integrated with their environments.[22]

A particularly important example of how computer games use—and extend—cinematic language is their implementation of a dynamic point of view. In driving and flying simulators and in combat games such as *Tekken 2* (Namco, 1994–), events like car crashes and knockdowns are automatically replayed from a different point of view. Other games such as the *Doom* series (Id Software, 1993–) and *Dungeon Keeper* (Bullfrog Productions, 1997) allow the user to switch between the point of view of the hero and a top-down bird's-eye view. Designers of online virtual worlds such as Active Worlds provide their users with similar capabilities. Nintendo goes even further by dedicating four buttons on its N64 joypad to controlling the view of the action. While playing Nintendo games such as *Super Mario 64* (Nintendo, 1996) the user can continuously adjust the position of the camera. Some Sony Playstation games such as *Tomb Raider* (Eidos, 1996) also use the buttons on the Playstation joypad for changing point of view. Some games such as *Myth: The Fallen Lords* (Bungie, 1997) use an AI engine (computer code that controls simulated "life" in the game, such as human characters that the player encounters) to automatically control the camera.

The incorporation of virtual/camera controls into the very hardware of game consoles is truly a historic event. Directing the virtual camera becomes as important as controlling the hero's actions. This fact is admitted by the game industry itself. Of the four key features of *Dungeon Keeper* advertized on its package, for instance, the first two concern control of the camera: "switch your perspective," "rotate your view," "take on your friend," "unveil hidden

22. Critical literature on computer games, and in particular, their visual language, remains slim. Useful facts on the history of computer games, descriptions of different genres, and interviews with designers can be found in Chris McGowan and Jim McCullaugh, *Entertainment in the Cyber Zone* (New York: Random House, 1995). Another useful source is J. C. Herz, *Joystick Nation: How Videogames Ate Our Quarters, Won Our Hearts, and Rewired Our Minds* (Boston: Little, Brown, 1997).



levels." In games such as this one, cinematic perception functions as the subject in its own right,[23] suggesting the return of "The New Vision" movement of the 1920s (Moholy-Nagy, Rodchenko, Vertov, and others), which foregrounded the new mobility of the photo and film camera, and made unconventional points of view a key part of its poetics.

The fact that computer games and virtual worlds continue to encode, step by step, the grammar of a kino-eye in software and in hardware is not an accident, but rather is consistent with the overall trajectory of the computerization of culture since the 1940s—the automation of all cultural operations. This automation gradually moves from basic to more complex operations: from image processing and spell checking to software-generated characters, 3-D worlds, and Web sites. A side effect of this automation is that once particular cultural codes are implemented in low-level software and hardware, they are no longer seen as choices but as unquestionable defaults. To take the automation of imaging as an example, in the early 1960s the newly emerging field of computer graphics incorporated a linear one-point perspective into 3-D software, and later directly into the hardware.[24] As a result, linear perspective became the default mode of vision in computer culture, whether we are speaking of computer animation, computer games, visualization, or VRML worlds. Now we are witnessing the next stage of this process—the translation of a cinematic grammar of points of view into software and hardware. As Hollywood cinematography is translated into algorithms and computer chips, its conventions become the default method of interacting with any data subjected to spatialization. (At SIGGRAPH '97 in Los Angeles, one of the presenters called for the incorporation of Hollywood-style editing in multi-user virtual worlds software. In such implementation, user interaction with other avatar(s) will be automatically rendered using classical Hollywood conventions for filming dialog.[25]) To use the terms of "The Virtual

23. *Dungeon Keeper* (Bullfrog Productions, 1997).

24. For a more detailed discussion of the history of computer imaging as gradual automation, see my articles "Mapping Space: Perspective, Radar, and Computer Graphics," and "Automation of Sight from Photography to Computer Vision."

25. Moses Ma's presentation on the panel "Putting a Human Face on Cyberspace: Designing Avatars and the Virtual Worlds They Live In," SIGGRAPH '97, 7 August 1997.

Cinematographer: A Paradigm for Automatic Real-Time Camera Control and Directing," a 1996 paper authored by Microsoft researchers, the goal of research is to encode "cinematographic expertise," translating "heuristics of filmmaking" into computer software and hardware.[26] Element by element, cinema is being poured into a computer: first, one-point linear perspective; next, the mobile camera and rectangular window; next, cinematography and editing conventions; and, of course, digital personas based on acting conventions borrowed from cinema, to be followed by make-up, set design, and the narrative structures themselves. Rather than being merely one cultural language among others, cinema is now becoming *the* cultural interface, a toolbox for all cultural communication, overtaking the printed word.

Cinema, the major cultural form of the twentieth century, has found a new life as the toolbox of the computer user. Cinematic means of perception, of connecting space and time, of representing human memory, thinking, and emotion have become a way of work and a way of life for millions in the computer age. Cinema's aesthetic strategies have become basic organizational principles of computer software. The window into a fictional world of a cinematic narrative has become a window into a datascape. In short, what was cinema is now the human-computer interface.

I will conclude this section by discussing a few artistic projects that, in different ways, offer alternatives to this trajectory—a trajectory that, again, involves the gradual translation of elements and techniques of cinematic perception and language into a de-contextualized set of tools to be used as an interface to any data. In the process of this translation, cinematic perception is divorced from its original material embodiment (camera, film stock), as well as from the historical context of its formation. If in cinema the camera functions as a material object, coexisting spatially and temporally with the world it is showing us, it has now becomes a set of abstract operations. The art projects that I discuss below refuse this separation of cinematic vision from the material world. They reunite perception and material reality by making the camera and what it records a part of the ontology of a virtual world. They

26. Li-wei He, Michael Cohen, and David Salesin, "The Virtual Cinematographer: A Paradigm for Automatic Real-Time Camera Control and Directing," SIGGRAPH '96 (http://research.microsoft.com/SIGGRAPH96/96/VirtualCinema.htm).



also refuse the universalization of cinematic vision by computer culture, which (just as postmodern visual culture in general) treats cinema as a toolbox, a set of "filters" that can be used to process any input. In contrast, each of these projects employs a unique cinematic strategy that has a specific relation to the particular virtual world it reveals to the user.

In *The Invisible Shape of Things Past*, Joachim Sauter and Dirk Lüsenbrink of the Berlin-based ART+COM collective created a truly innovative cultural interface for accessing historical data about Berlin's history.[27] The interface de-virtualizes cinema, so to speak, by putting the records of cinematic vision back into their historical and material context. As the user navigates through a 3-D model of Berlin, she comes across elongated shapes lying on city streets. These shapes, which the authors call "filmobjects," correspond to documentary footage recorded at corresponding points in the city. To create each shape, the original footage is digitized and the frames are stacked one after another in depth, with the original camera parameters determining the exact shape. The user can view the footage by clicking on the first frame. As the frames are displayed one after another, the shape becomes correspondingly thinner.

In following the general trend of computer culture toward spatialization of every cultural experience, this cultural interface spatializes time, representing it as a shape in a 3-D space. This shape can be thought of as a book, with individual frames stacked one after another like book pages. The trajectory through time and space followed by a camera becomes a book to be read, page by page. The records of the camera's vision become material objects, sharing space with the material reality that gave rise to this vision. Cinema is solidified. This project, then, can be also understood as a virtual monument to cinema. The (virtual) shapes situated around the (virtual) city remind us of the era when cinema was the defining form of cultural expression—as opposed to a toolbox for data retrieval and use.

Hungarian-born artist Tamás Waliczky openly refuses the default mode of vision imposed by computer software—one-point linear perspective. Each of his computer-animated films *The Garden* (1992), *The Forest* (1993) and *The Way* (1994) utilizes a particular perspectival system: a water-drop

27. See http://www.artcom.de/projects/invisible_shape/welcome.en.

perspective in *The Garden,* a cylindrical perspective in *The Forest,* and a reverse perspective in *The Way.* Working with computer programmers, the artist created custom-made 3-D software to implement these perspectival systems. Each of the systems has an inherent relationship to the subject of the film in which it is used. In *The Garden,* the subject is the perception of a small child, for whom the world does not yet have an objective existence. In *The Forest,* the mental trauma of emigration is translated into the endless roaming of a camera through the forest, which is actually just a set of transparent cylinders. Finally, in *The Way,* the self-sufficiency and isolation of a Western subject are conveyed by the use of a reverse perspective.

In Waliczky's films the camera and the world are made into a single whole, whereas in *The Invisible Shape of Things Past* the records of the camera are placed back into the world. Rather than simply subjecting his virtual worlds to different types of perspectival projection, Waliczky modified the spatial structure of the worlds themselves. In *The Garden,* a child playing in a garden becomes the center of the world; as she moves around, the actual geometry of all the objects around her is transformed, with objects becoming bigger as she gets closer to them. To create *The Forest,* a number of cylinders were placed inside each other, each cylinder mapped with a picture of a tree, repeated a number of times. In the film, we see a camera moving through this endless static forest in a complex spatial trajectory—but this is an illusion. In reality, the camera does move, but the architecture of the world is constantly changing as well, because each cylinder is rotating at its own speed. As a result, the world and our perception of it are fused together.

### HCI: Representation versus Control

The development of the human-computer interface, until recently, has had little to do with the distribution of cultural objects. Following some of the main applications from the 1940s until the early 1980s, when the current generation of the GUI was developed and reached the mass market together with the rise of the PC, we can list the most significant: real-time control of weapons and weapon systems; scientific simulation; computer-aided design; and finally, office work with the secretary functioning as prototypical computer user—filing documents in folders, emptying the trash can, creating and editing documents ("word processing"). Today, as the computer is beginning to host very different applications for access and manipulation of cultural data and cultural experiences, their interfaces still rely on old metaphors and



action grammars. Cultural interfaces predictably use elements of a general-purpose HCI such as scrollable windows containing text and other data types, hierarchical menus, dialogue boxes, and command-line input. For instance, a typical "art collection" CD-ROM tries to recreate "the museum experience" by presenting a navigable 3-D rendering of a museum space, while still resorting to hierarchical menus that allow the user to switch between different museum collections. Even in the case of *The Invisible Shape of Things Past,* which uses a unique interface solution of "filmobjects" not directly traceable to either old cultural forms or general-purpose HCI, the designers still rely on HCI convention in the use of a pull-down menu to switch between different maps of Berlin.

In their important study of new media, *Remediation,* Jay David Bolter and Richard Grusin define *medium* as "that which remediates."[28] In contrast to a modernist view that aims to define the essential properties of every medium, Bolter and Grusin propose that all media work by "remediating," that is, translating, refashioning, and reforming other media, both on the level of content and form. If we think of the human-computer interface as another medium, its history and present development definitely fit this thesis. The history of the human-computer interface is that of borrowing and reformulating, or, to use new media lingo, reformatting other media, both past and present—the printed page, film, television. But along with borrowing the conventions of most other media and eclectically combining them together, HCI designers also heavily borrow "conventions" of the human-made physical environment, beginning with Macintosh's use of the desktop metaphor. And, more than any medium before it, HCI is like a chameleon that keeps changing its appearance, responding to how computers are used in any given period. For instance, if in the 1970s the designers at Xerox PARC modeled the first GUI on the office desk because they imagined that the computer they were designing would be used in the office, in the 1990s the primary use of computers as media-access machines led to the borrowing of interfaces of already familiar media devices such as the VCR or audio CD player controls.

---

28. Jay David Bolter and Richard Grusin, *Remediation: Understanding New Media* (Cambridge, Mass.: MIT Press, 1999), 19.

In general, cultural interfaces of the 1990s try to walk an uneasy path between the richness of control provided in general-purpose HCI and the "immersive" experience of traditional cultural objects such as books and movies. Modern general-purpose HCI, be it the MAC OS, Windows, or UNIX, allow their users to perform complex and detailed actions on computer data: acquire information about an object, copy it, move it to another location, change the way data is displayed, etc. In contrast, a conventional book or a film positions the user inside an imaginary universe whose structure is fixed by the author. Cultural interfaces attempt to mediate between these two fundamentally different and ultimately incompatible approaches.

As an example, consider how cultural interfaces conceptualize the computer screen. If a general-purpose HCI clearly identifies to the user that certain objects can be acted on while others cannot (icons representing files but not the desktop itself), cultural interfaces typically hide the hyperlinks within a continuous representational field. (This technique was already so widely accepted by the 1990s that the designers of HTML offered it early on to users by implementing the "imagemap" feature.) The field can be a two-dimensional collage of different images, a mixture of representational elements and abstract textures, or a single image of a space such as a city street or a landscape. By trial and error, clicking all over the field, the user discovers that some parts of this field are hyperlinks. This concept of a screen combines two distinct pictorial conventions—the older Western tradition of pictorial illusionism in which a screen functions as a window into a virtual space, something for the viewer to look into but not act upon; and the more recent convention of graphical human-computer interfaces that divides the computer screen into a set of controls with clearly delineated functions, thereby essentially treating it as a virtual instrument panel. As a result, the computer screen becomes a battlefield for a number of incompatible definitions—depth and surface, opaqueness and transparency, image as illusionary space and image as instrument for action.

The computer screen also functions both as a window into an illusionary space and as a flat surface carrying text labels and graphical icons. We can relate this to a similar understanding of a pictorial surface in the Dutch art of the seventeenth century. In her classic study *The Art of Describing*, art historian Svetlana Alpers discusses how Dutch painting of the period functioned as both map and picture, combining different kinds of information and knowledge of the world.[29]

Here is another example of how cultural interfaces try to find a middle ground between the conventions of general-purpose HCI and the conventions of traditional cultural forms. Again we encounter tension and struggle—in this case, between standardization and originality. One of the main principles of modern HCI is the consistency principle. It dictates that menus, icons, dialogue boxes, and other interface elements should be the same in different applications. The user knows that every application will contain a "file" menu, or that if she encounters an icon that looks like a magnifying glass, it can be used to zoom on documents. In contrast, modern culture (including its "postmodern" stage) stresses originality: Every cultural object is supposed to be different from the rest, and if it is quoting other objects, these quotes have to be defined as such. Cultural interfaces try to accommodate both the demand for consistency and the demand for originality. Most of them contain the same set of interface elements with standard semantics, such as "home," "forward," and "backward" icons. But because every Web site and CD-ROM strives to have its own distinct design, these elements are always designed differently from one product to the next. For instance, many games such as *War Craft II* (Blizzard Entertainment, 1996) and *Dungeon Keeper* give their icons a "historical" look consistent with the mood of the imaginary universe portrayed in the game.

The language of cultural interfaces is a hybrid. It is a strange, often awkward mix between the conventions of traditional cultural forms and the conventions of HCI—between an immersive environment and a set of controls, between standardization and originality. Cultural interfaces try to balance the concept of a surface in painting, photography, cinema, and the printed page as something to be looked at, glanced at, read, but always from some distance, without interfering with it, with the concept of the surface in a computer interface as a virtual control panel, similar to the control panel on

29. See Svetlana Alpers, *The Art of Describing: Dutch Art in the Seventeenth Century* (Chicago: University of Chicago Press, 1983). See particularly the chapter "Mapping Impulse."

a car, plane, or any other complex machine.[30] Finally, on yet another level, the traditions of the printed word and of cinema also compete between themselves. One wants the computer screen to be a dense and flat information surface, whereas the other insists that it become a window into a virtual space.

To see that this hybrid language of the cultural interfaces of the 1990s represents only one historical possibility, consider a very different scenario. Potentially, cultural interfaces could completely rely on already existing metaphors and action grammars of a standard HCI, or, at least, rely on them much more than they actually do. They do not have to "dress up" HCI with custom icons and buttons, or hide links within images, or organize the information as a series of pages or a 3-D environment. For instance, texts can be presented simply as files inside a directory rather than as a set of pages connected by custom-designed icons. This strategy of using standard HCI to present cultural objects is encountered quite rarely. In fact, I am aware of only one project that seems to use it completely consciously, as though by choice rather than by necessity—a CD-ROM by Gerald Van Der Kaap entitled *BlindRom V.0.9.* (Netherlands, 1993). The CD-ROM includes a standard-looking folder named "Blind Letter." Inside the folder there are a large number of text files. You do not have to learn yet another cultural interface, search for hyperlinks hidden in images, or navigate through a 3-D environment. Reading these files requires simply opening them in standard Macintosh SimpleText, one by one. This simple technique works very well. Rather than distracting the user from experiencing the work, the computer interface becomes part and parcel of the work. Opening these files, I felt that I was in the presence of a new literary form for a new medium, perhaps the real medium of a computer—its interface.

As the examples here illustrate, cultural interfaces try to create their own language rather than simply using the general-purpose HCI. In doing so, these

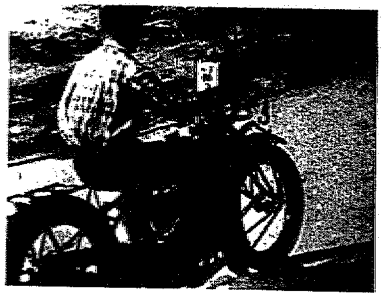
30. This historical connection is illustrated by popular flight simulator games in which the computer screen is used to simulate the control panel of a plane, that is, the very type of object from which computer interfaces have developed. The conceptual origin of the modern GUI in a traditional instrument panel can be seen even more clearly in the first graphical computer interfaces of the late 1960s and early 1970s, which used tiled windows. The first tiled window interface was demonstrated by Douglas Engelbart in 1968.



interfaces try to negotiate between metaphors and ways of controlling a computer developed in HCI, and the conventions of more traditional cultural forms. Indeed, neither extreme is ultimately satisfactory by itself. It is one thing to use a computer to control weapons or analyze statistical data, it is another to use it to represent cultural memories, values, and experiences. Interfaces developed for the computer in the role of calculator, control mechanism, or communication device are not necessarily suitable for a computer playing the role of cultural machine. Conversely, if we simply mimic the existing conventions of older cultural forms such as the printed word and cinema, we will not take advantage of all the new capacities offered by the computer: its flexibility in displaying and manipulating data, interactive control by the user, ability to run simulations, etc.

Today the language of cultural interfaces is in its early stage, as was the language of cinema a hundred years ago. We do not know what the final result will be, or even if it will ever completely stabilize. Both the printed word and cinema eventually achieved stable forms that underwent little change for long periods of time, in part because of the material investments in their means of production and distribution. Given that computer language is implemented in software, potentially it could keep changing forever. But there is one thing we can be sure of. We are witnessing the emergence of a new cultural metalanguage, something that will be at least as significant as the printed word and cinema before it.

# The Screen and the User



Contemporary human-computer interfaces offer radical new possibilities for art and communication. Virtual reality allows us to travel through nonexistent three-dimensional spaces. A computer monitor connected to a network becomes a window through which we can enter places thousands of miles away. Finally, with the help of a mouse or a video camera, a computer can be transformed into an intelligent being capable of engaging us in dialogue.

VR, telepresence, and interactivity are made possible by the recent technology of the digital computer. However, they are made real by a much older technology—the screen. It is by looking at a screen—a flat, rectangular surface positioned at some distance from the eyes—that the user experiences the illusion of navigating through virtual spaces, of being physically present somewhere else or of being hailed by the computer itself. If computers have become a common presence in our culture only in the last decade, the screen, on the other hand, has been used to present visual information for centuries—from Renaissance painting to twentieth-century cinema.

Today, coupled with the computer, the screen is rapidly becoming the main means of accessing any kind of information, be it still images, moving images, or text. We are already using it to read the daily newspaper; to watch movies; to communicate with co-workers, relatives, and friends; and, most important, to work. We may debate whether our society is a society of spectacle or of simulation, but, undoubtedly, it is a society of the screen. What are the different stages of the screen's history? What are the relationships between the physical space where the viewer is located, her body, and the screen



space? What are the ways in which computer displays both continue and challenge the tradition of the screen?[31]

## A Screen's Genealogy

Let us start with the definition of a screen. The visual culture of the modern period, from painting to cinema, is characterized by an intriguing phenomenon—the existence of *another* virtual space, another three-dimensional world enclosed by a frame and situated inside our normal space. The frame separates two absolutely different spaces that somehow coexist. This phenomenon is what defines the screen in the most general sense, or, as I will call it, the "classical screen."

What are the properties of a classical screen? It is a flat, rectangular surface. It is intended for frontal viewing—as opposed to a panorama for instance. It exists in our normal space, the space of our body, and acts as a window into another space. This other space, the space of representation, typically has a scale different from the scale of our normal space. Defined in this way, a screen describes equally well a Renaissance painting (recall Alberti's formulation referred to above) and a modern computer display. Even proportions have not changed in five centuries; they are similar for a typical fifteenth-century painting, a film screen, and a computer screen. In this respect it is not accidental that the very names of the two main formats of

---

31. My analysis here focuses on the continuities between the computer screen and preceding representational conventions and technologies. For alternative readings that take up the differences between the two, see the excellent articles by Vivian Sobchack, "Nostalgia for a Digital Object: Regrets on the Quickening of QuickTime," in *Millennium Film Journal* 4–23, No. 34 (Fall 1999) and Norman Bryson, "Summer 1999 at TATE," available from Tate Gallery, 413 West 14th Street, New York City. Bryson writes: "Though the [computer] screen is able to present a scenographic depth, it is obviously unlike the Albertian or Renaissance window; its surface never vanishes before the imaginary depths behind it, it never truly opens into depth. But the PC screen does not behave like the modernist image, either. It cannot foreground the materiality of the surface (of pigments on canvas) since it has no materiality to speak of, other than the play of shifting light." Both Sobchack and Bryson stress the difference between the traditional image frame and the multiple windows of a computer screen. "Basically," writes Bryson, "the whole order of the frame is abolished, replaced by the order of superimposition or tiling."

computer displays point to two genres of painting: A horizontal format is referred to as "landscape mode," whereas the vertical format is referred to as "portrait mode."

A hundred years ago a new type of screen, which I will call the "dynamic screen," became popular. This new type retains all the properties of a classical screen while adding something new: It can display an image changing over time. This is the screen of cinema, television, video. The dynamic screen also brings with it a certain relationship between the image and the spectator—a certain *viewing regime,* so to speak. This relationship is already implicit in the classical screen, but now it fully surfaces. A screen's image strives for complete illusion and visual plenitude, while the viewer is asked to suspend disbelief and to identify with the image. Although the screen in reality is only a window of limited dimensions positioned inside the physical space of the viewer, the viewer is expected to concentrate completely on what she sees in this window, focusing her attention on the representation and disregarding the physical space outside. This viewing regime is made possible by the fact that the singular image, whether a painting, movie screen, or television screen, completely fills the screen. This is why we are so annoyed in a movie theater when the projected image does not precisely coincide with the screen's boundaries: It disrupts the illusion, making us conscious of what exists outside the representation.[32]

Rather than being a neutral medium of presenting information, the screen is aggressive. It functions to filter, to *screen out,* to take over, rendering nonexistent whatever is outside its frame. Of course, the degree of this filtering varies between cinema viewing and television viewing. In cinema viewing, the viewer is asked to merge completely with the screen's space. In television viewing (as it was practiced in the twentieth century), the screen is smaller, lights are on, conversation between viewers is allowed, and the act of viewing is often integrated with other daily activities. Still, overall this viewing regime has remained stable—until recently.

32. The degree to which a frame that acts as a boundary between the two spaces is emphasized seems to be proportional to the degree of identification expected from the viewer. Thus in cinema, where the identification is most intense, the frame as a separate object does not exist at all—the screen simply ends at its boundaries—whereas both in painting and television the framing is much more pronounced.



This stability has been challenged by the arrival of the computer screen. On the one hand, rather than showing a single image, a computer screen typically displays a number of coexisting windows. Indeed, the coexistence of a number of overlapping windows is a fundamental principle of the modern GUI. No single window completely dominates the viewer's attention. In this sense, the possibility of simultaneously observing a few images that coexist within one screen can be compared with the phenomenon of zapping—the quick switching of television channels that allows the viewer to follow more than program.[33] In both instances, the viewer no longer concentrates on a single image. (Some television sets enable a second channel to be watched within a smaller window positioned in a corner of the main screen. Perhaps future TV sets will adopt the window metaphor of a computer.) A window interface has more to do with modern graphic design, which treats a page as a collection of different but equally important blocks of data such as text, images, and graphic elements, than with the cinematic screen.

On the other hand, with VR, the screen disappears altogether. VR typically uses a head-mounted display whose images completely fill the viewer's visual field. No longer is the viewer looking at a rectangular, flat surface from a certain distance, a window into another space. Now she is fully situated within this other space. Or, more precisely, we can say that the two spaces—the real, physical space and the virtual, simulated space—coincide. The virtual space, previously confined to a painting or a movie screen, now completely encompasses the real space. Frontality, rectangular surface, difference in scale are all gone. The screen has vanished.

Both situations—window interface and VR—disrupt the viewing regime that characterizes the historical period of the dynamic screen. This regime, based on an identification of viewer and screen image, reached its culmination in the cinema, which goes to an extreme to enable this identification (the bigness of the screen, the darkness of the surrounding space).

Thus, the era of the dynamic screen that began with cinema is now ending. And it is this disappearance of the screen—its splitting into many windows in window interface, its complete takeover of the visual field in

33. Here I agree with the parallel suggested by Anatoly Prokhorov between window interface and montage in cinema.

VR—that allows us today to recognize it as a cultural category and begin to trace its history.

The origins of the cinema's screen are well known. We can trace its emergence to the popular spectacles and entertainments of the eighteenth and nineteenth centuries: magic lantern shows, phantasmagoria, eidophusikon, panorama, diorama, zoopraxiscope shows, and so on. The public was ready for cinema, and when it finally appeared, it was a huge public event. Not by accident, the "invention" of cinema was claimed by at least a dozen individuals from a half-dozen countries.[34]

The origin of the computer screen is a different story. It appears in the middle of this century, but it does not become a public presence until much later; and its history has not yet been written. Both of these facts are related to the context in which it emerged: As with all the other elements of modern human-computer interface, the computer screen was developed for military use. Its history has to do not with public entertainment but with military surveillance.

The history of modern surveillance technologies begins with photography. With the advent of photography came an interest in using it for aerial surveillance. Félix Tournachon Nadar, one of the most eminent photographers of the nineteenth century, succeeded in exposing a photographic plate at 262 feet over Bièvre, France in 1858. He was soon approached by the French Army to attempt photo reconnaissance but rejected the offer. In 1882, unmanned photo balloons were already in the air; a little later, they were joined by photo rockets both in France and in Germany. The only innovation of World War I was to combine aerial cameras with a superior flying platform—the airplane.[35]

Radar became the next major surveillance technology. Massively employed in World War II, it provided important advantages over photography. Previously, military commanders had to wait until pilots returned from surveillance missions and film was developed. The inevitable delay between time of surveillance and delivery of the finished image limited photography's usefulness because by the time a photograph was produced, enemy positions

34. For these origins see, for instance, C. W. Ceram, *Archeology of the Cinema* (New York: Harcourt Brace and World, 1965).

35. Beaumont Newhall, *Airborne Camera* (New York: Hastings House, 1969).



could have changed. However, with radar, imaging became instantaneous, and this delay was eliminated. The effectiveness of radar had to do with a new means of displaying an image—a new type of screen.

Consider the imaging technologies of photography and film. The photographic image is a permanent imprint corresponding to a single referent—whatever is in front of the lens when the photograph is taken. It also corresponds to a limited time of observation—the time of exposure. Film is based on the same principles. A film sequence, composed of a number of still images, represents the sum of referents and the sum of exposure times of these individual images. In either case, the image is fixed once and for all. Therefore the screen can only show past events.

With radar, we see for the first time the mass employment (television is founded on the same principle but its mass employment comes later) of a fundamentally new type of screen, a screen that gradually comes to dominate modern visual culture—video monitor, computer screen, instrument display. What is new about such a screen is that its image can change in real time, reflecting changes in the referent, whether the position of an object in space (radar), any alteration in visible reality (live video) or changing data in the computer's memory (computer screen). The image can be continually updated *in real time.* This is the third type of screen after classic and dynamic—the screen of real time.

The radar screen changes, tracking the referent. But while it appears that the element of time delay, always present in the technologies of military surveillance, is eliminated, in fact, time enters the real-time screen in a new way. In older, photographic technologies, all parts of an image are exposed simultaneously, whereas now the image is produced through sequential scanning—circular in the case of radar, horizontal in the case of television. Therefore, the different parts of the image correspond to different moments in time. In this respect, a radar image is more similar to an audio record, since consecutive moments in time become circular tracks on a surface.[36]

36. This is more than a conceptual similarity. In the late 1920s, John H. Baird invented "phonovision," the first method for the recording and playback of a television signal. The signal was recorded on Edison's phonograph record by a process very similar to that of making an audio recording. Baird named his recording machine the "phonoscope." Albert Abramson, *Electronic Motion Pictures* (University of California Press, 1955), 41–42.

What this means is that the image, in a traditional sense, no longer exists! And it is only by habit that we still refer to what we see on the real-time screen as "images." It is only because the scanning is fast enough and because, sometimes, the referent remains static, that we see what looks like a static image. Yet, such an image is no longer the norm, but the exception of a more general, new kind of representation for which we do not yet have a term.

The principles and technology of radar were worked out independently by scientists in the United States, England, France, and Germany during the 1930s. After the beginning of the War, however, only the U.S. had the resources necessary to continue radar development. In 1940, at MIT, a team of scientists was assembled to work in the Radiation Laboratory, or the "Rad Lab," as it came to be called. The purpose of the lab was radar research and production. By 1943, the "Rad Lab" occupied 115 acres of floor space; it had the largest telephone switchboard in Cambridge and employed four thousand people.[37]

Next to photography, radar provided a superior way to gather information about enemy locations. In fact, it provided too much information, more information than one person could deal with. Historical footage from the early days of the war shows a central command room with a large, table-size map of Britain.[38] Small pieces of cardboard in the form of planes are positioned on the map to show the locations of actual German bombers. A few senior officers scrutinize the map. Meanwhile, women in army uniforms constantly change the location of the cardboard pieces by moving them with long sticks as information is transmitted from dozens of radar stations.[39]

Was there a more effective way to process and display information gathered by radar? The computer screen, as well as most other key principles and technologies of the modern human-computer interface—interactive control, algorithms for 3-D wireframe graphics, bit-mapped graphics—was developed as a way of solving this problem.

The research again took place at MIT. The Radiation Laboratory was dismantled after the end of the war, but soon the Air Force created another

37. *Echoes of War* (Boston: WGBH Boston, 1989), videotape.

38. Ibid.

39. Ibid.



secret laboratory in its place—Lincoln Laboratory. The purpose of Lincoln Laboratory was to work on human factors and new display technologies for SAGE—"Semi-Automatic Ground Environment," a command center to control the U.S. air defenses established in the mid-1950s.[40] Historian of computer technology Paul Edwards writes that SAGE's job "was to link together radar installations around the USA's perimeter, analyze and interpret their signals, and direct manned interceptor jets toward the incoming bee. It was to be a total system, one whose 'human components' were fully integrated into the mechanized circuit of detection, decision and response."[41]

The creation of SAGE and the development of an interactive human-computer interface were largely the result of a particular military doctrine. In the 1950s, the American military thought that a Soviet attack on the U.S. would entail sending a large number of bombers simultaneously. Therefore, it seemed necessary to create a center that could receive information from all U.S. radar stations, track the large number of enemy bombers, and coordinate a counterattack. The computer screen and other components of the modern human-computer interface owe their existence to this particular military idea. (As someone who was born in the Soviet Union and now works on the history of new media in the United States, I find this bit of history truly fascinating.)

An early version of the center was called "the Cape Cod network," since it received information from radars situated along the coast of New England. The center operated right out of the Barta Building on the MIT campus. Each of eighty-two Air Force officers monitored his own computer display, which showed the outline of the New England Coast and the location of key radars. Whenever an officer noticed a dot indicating a moving plane, he

40. On SAGE, see the excellent social history of early computing by Paul Edwards, *The Closed World: Computers and the Politics of Discourse in Cold War America* (Cambridge, Mass.: MIT Press, 1996). For a shorter summary of his argument, see Paul Edwards, "The Closed World: Systems Discourse, Military Policy and Post–World War II U.S. Historical Consciousness," in *Cyborg Worlds: The Military Information Society*, eds. Les Levidow and Kevin Robins (London: Free Association Books, 1989). See also Howard Rheingold, *Virtual Reality* (New York: Simon and Schuster, 1991), 68–93.

41. Edwards, "The Closed World" (1989), 142.

would tell the computer to follow the plane. To do this, the officer simply had to touch the dot with a special "light pen."[42]

Thus, the SAGE system contained all the main elements of the modern human-computer interface. The light pen, designed in 1949, can be considered a precursor of the contemporary mouse. More importantly, at SAGE, the screen came to be used not only to display information in real time, as in radar and television, but also to give commands to the computer. Rather than acting solely as a means of displaying an image of reality, the screen became a vehicle for directly affecting reality.

Using the technology developed for SAGE, Lincoln researchers created a number of computer graphics programs that relied on the screen as a means of inputting and outputting information from a computer. These included programs for displaying brain waves (1957), simulating planet and gravitational activity (1960), and creating 2-D drawings (1958).[43] The most well-known of these programs was "Sketchpad." Designed in 1962 by Ivan Sutherland, a graduate student supervised by Claude Shannon, it widely publicized the idea of interactive computer graphics. With Sketchpad, a human operator could create graphics directly on a computer screen by touching the screen with a light pen. Sketchpad exemplified a new paradigm of interacting with computers: By changing something on the screen, the operator changed something in the computer's memory. The real-time screen became interactive.

This, in short, is the history of the birth of the computer screen. But even before the computer screen became widely used, a new paradigm emerged—the simulation of an interactive three-dimensional environment without a screen. In 1966, Ivan Sutherland and his colleagues began research on the prototype of VR. The work was cosponsored by the Advanced Research Projects Agency (ARPA) and the Office of Naval Research.[44]

"The fundamental idea behind the three-dimensional display is to present the user with a perspective image which changes as he moves," wrote

42. "Retrospectives II: The Early Years in Computer Graphics at MIT, Lincoln Lab, and Harvard," in *SIGGRAPH '89 Panel Proceedings* (New York: The Association for Computing Machinery, 1989), 22–24.

43. Ibid., 42–54.

44. Rheingold, *Virtual Reality*, 105.



Sutherland in 1968.[45] The computer tracked the position of the viewer's head and adjusted the perspective of the computer graphic image accordingly. The display itself consisted of two six-inch-long monitors mounted next to the temples. They projected an image that appeared superimposed over the viewer's field of vision.

The screen disappeared. It had completely taken over the visual field.

## The Screen and the Body

I have presented one possible genealogy of the modern computer screen. In my genealogy, the computer screen represents an interactive type, a subtype of the real-time type, which is a subtype of the dynamic type, which is a subtype of the classical type. My discussion of these types relied on two ideas. First, the idea of temporality—the classical screen displays a static, permanent image; the dynamic screen displays a moving image of the past; and finally, the real-time screen shows the present. Second, the relationship between the space of the viewer and the space of representation (I defined the screen as a window into the space of representation that itself exists in our normal space).

Let us now look at the screen's history from another angle—the relationship between the screen and the body of the viewer. This is how Roland Barthes describes the screen in "Diderot, Brecht, Eisenstein," written in 1973:

> Representation is not defined directly by imitation: even if one gets rid of notions of the "real," of the "vraisemblable," of the "copy," there will still be representation for as long as a subject (author, reader, spectator or voyeur) casts his *gaze* towards a horizon on which he cuts out a base of a triangle, his eye (or his mind) forming the apex. The "Organon of Representation" (which is today becoming possible to write because there are intimations of *something else*) will have as its dual foundation the sovereignty of the act of cutting out [*découpage*] and the unity of the subject of action. . . . The scene, the picture, the shot, the cut-out rectangle, here we have the very *condition* that allows us to conceive theater, painting, cinema, literature, all those arts, that is, other than music and which could be called *dioptric arts*.[46]

45. Quoted in ibid., 104.

46. Roland Barthes, "Diderot, Brecht, Eisenstein," in *Image/Music/Text*, trans. Stephen Heath (New York: Farrar, Straus, and Giroux, 1977), 69–70.

For Barthes, the screen becomes an all-encompassing concept that covers the functioning of even non-visual representation (literature), although he does make an appeal to a particular visual model of linear perspective. At any rate, his concept encompasses all the types of representational apparatuses I have discussed: painting, film, television, radar, and computer display. In each of these, reality is cut by the rectangle of a screen: "a pure cut-out segment with clearly defined edges, irreversible and incorruptible; everything that surrounds it is banished into nothingness, remains unnamed, while everything that it admits within its field is promoted into essence, into light, into view."[47] This act of cutting reality into a sign and nothingness simultaneously doubles the viewing subject, who now exists in two spaces: the familiar physical space of her real body and the virtual space of an image within the screen. This split comes to the surface with VR, but it already exists in painting and other *dioptric arts*.

What is the price the subject pays for the mastery of the world, focused and unified by the screen?

*The Draughtsman's Contract*, a 1982 film by Peter Greenaway, concerns an architectural draftsman hired to produce a set of drawings of a country house. The draughtsman employs a simple drawing tool consisting of a square grid. Throughout the film, we repeatedly see the draughtsman's face through the grid, which looks like prison bars. It is as if the subject who attempts to catch the world, immobilizing and fixing it within the representational apparatus (here, perspectival drawing), is trapped by the apparatus himself. The subject is imprisoned.

I take this image as a metaphor for what appears to be a general tendency of the Western screen-based representational apparatus. In this tradition, the body must be fixedin space if the viewer is to see the image at all. From Renaissance monocular perspective to modern cinema, from Kepler's camera obscura to nineteenth-century camera lucida, the body has to remain still.[48]

47. Ibid.

48. Although in the following I discuss the immobility of the subject of a screen in the context of the history of representation, we can also relate this condition to the history of communication. In ancient Greece, communication was understood as an oral dialogue between people. It was also assumed that physical movement stimulated dialogue and the process of thinking. Aristotle and his pupils walked around while discussing philosophical problems. In



The imprisonment of the body takes place on both the conceptual and literal levels; both kinds of imprisonment already appear with the first screen apparatus, Alberti's perspectival window, which, according to many interpreters of linear perspective, presents the world as seen by a singular eye—static, unblinking, and fixated. As described by Norman Bryson, perspective "followed the logic of the Gaze rather than the Glance, thus producing a visual take that was eternalized, reduced to one 'point of view' and disembodied."[49] Bryson argues that "the gaze of the painter arrests the flux of phenomena, contemplates the visual field from a vantage point outside the mobility of duration, in an eternal moment of disclosed presence."[50] Correspondingly, the world, as seen by this immobile, static, and atemporal Gaze, which belongs more to a statue than a living body, becomes equally immobile, reified, fixated, cold and dead. Referring to Dürer's famous print of a draftsman drawing a nude through a screen of perspectival threads, Martin Jay notes that "a reifying male look" turns "its targets into stone"; consequently, "the marmoreal nude is drained of its capacity to arouse desire."[51] Similarly, John Berger compares Alberti's window to "a safe let into a wall, a safe into which the visible has been deposited."[52] And in *The Draughtsman's Contract*, the draughtsman, time and again, tries to eliminate all motion, any sign of life, from the scenes he is rendering.

With perspectival machines, the imprisonment of the subject also happens in a literal sense. From the onset of the adaptation of perspective, artists and draftsmen attempted to aid the laborious manual process of creating perspectival images, and between the sixteenth and nineteenth centuries various "perspectival machines" were constructed.[53] By the first decades of the

the Middle Ages, a shift occured from dialogue between subjects to communication between a subject and an information storage device, that is, a book. A medieval book chained to a table can be considered a precursor to the screen that "fixes" its subject in space.

49. As summarized by Martin Jay, "Scopic Regimes of Modernity," in *Vision and Visuality*, ed. Hal Foster (Seattle: Bay Press, 1988), 7.

50. Quoted in ibid., 7.

51. Ibid., 8.

52. Quoted in ibid., 9.

53. For a survey of perspectival instruments, see Martin Kemp, *The Science of Art* (New Haven: Yale University Press, 1990), 167–220.

sixteenth century, Dürer had described a number of such machines.[54] Many varieties were invented, but regardless of the type, the artist had to remain immobile throughout the process of drawing.

Along with perspectival machines, a whole range of optical apparatuses was in use, particularly for depicting landscapes and conducting topographical surveys. The most popular optical apparatus was the camera obscura.[55] *Camera obscura* literally means "dark chamber," and was founded on the premise that if rays of light from an object or a scene pass through a small aperture, they will cross and reemerge on the other side to form an image on a screen. In order for the image to become visible, however, "it is necessary that the screen be placed in a chamber in which light levels are considerably lower than those around the object."[56] Thus, in one of the earliest depictions of the camera obscura, in Kircher's *Ars magna Lucis et umbrae* (Rome, 1649), we see the subject enjoying the image inside a tiny room, oblivious to the fact that he has had to imprison himself inside this "dark chamber" in order to see the image on the screen.

Later, a smaller tent-type camera obscura—a movable prison, so to speak—became popular. It consisted of a small tent mounted on a tripod, with a revolving reflector and lens at its apex. Having positioned himself inside the tent, which provided the necessary darkness, the draftsman would then spend hours meticulously tracing the image projected by the lens.

Early photography continued the trend toward the imprisonment of the subject and the object of representation. During photography's first decades, exposure times were quite long. The daguerreotype process, for instance, required exposures of four to seven minutes in the sun and from twelve to sixty minutes in diffused light. So, similar to the drawings produced with the help of the camera obscura, which depicted reality as static and immobile, early photographs represented the world as stable, eternal, unshakable. And when photography ventured to represent living things, they had to be immobilized. Thus, portrait studios universally employed various clamps to assure the steadiness of the sitter throughout the lengthy time of exposure. Reminiscent of torture instruments, the iron clamps firmly held the subject in

54. Ibid., 171–172.
55. Ibid., 200.
56. Ibid.



place—a subject who voluntarily became the prisoner of the machine in order to see her own image.[57]

Toward the end of the nineteenth century, the petrified world of the photographic image was shattered by the dynamic screen of the cinema. In "The Work of Art in the Age of Mechanical Reproduction," Walter Benjamin expressed his fascination with the new mobility of the visible: "Our taverns and our metropolitan streets, our offices and furnished rooms, our railroad stations and our factories appeared to have us locked up hopelessly. When came the film and burst this prison-world asunder by the dynamite of the tenth of a second, so that now, in the midst of its far-flung ruins and debris, we calmly and adventurously go traveling."[58]

The cinema screen enabled audiences to take a journey through different spaces without leaving their seats; in the words of film historian Anne Friedberg, it created "a mobilized virtual gaze."[59] However, the cost of this virtual mobility was a new, institutionalized immobility of the spectator. All around the world large prisons were constructed that could hold hundreds of prisoners—movie houses. The prisoners could neither talk to one another nor move from seat to seat. While they were taken on virtual journeys, their bodies remained still in the darkness of collective cameras obscura.

The formation of this viewing regime took place in parallel with the shift from what film theorists call "primitive" to "classical" film language.[60] An important part of this shift, which took place in the 1910s, was the new functioning of the virtual space represented on the screen. During the "primitive" period, the space of the film theater and the screen space were clearly separated, much like in theater or vaudeville. Viewers were free to interact, come and go, and maintain a psychological distance from the virtual world of the cinematic narrative. In contrast, classical film addressed each viewer as a separate individual and positioned him inside its virtual world

57. Anesthesiology emerges approximately at the same time.
58. Walter Benjamin, "The Work of Art in the Age of Mechanical Reproduction," in *Illuminations*, ed. Hannah Arendt (New York: Schocken Books, 1969), 238.
59. Anne Friedberg, *Window Shopping: Cinema and the Postmodern* (Berkeley: University of California Press, 1993), 2.
60. See, for instance, David Bordwell, Janet Steiger, and Kristin Thompson, *The Classical Hollywood Cinema* (New York: Columbia University Press, 1985).

narrative. As noted by a contemporary in 1913, "[spectators] should be put in the position of being a 'knot hole in the fence' at every stage in the play."[61] If "primitive cinema keeps the spectator looking across a void in a separate space,"[62] classical cinema positions the spectator in terms of the best viewpoint of each shot, inside the virtual space.

This situation is usually conceptualized in terms of the spectator's identification with the camera eye. The body of the spectator remains in her seat while her eye is coupled with a mobile camera. However, it is also possible to conceptualize this differently. We can imagine that the camera does not, in fact, move at all, but rather remains stationary, coinciding with the spectator's eyes. Instead, it is the virtual space as a whole that changes its position with each shot. Using the contemporary vocabulary of computer graphics, we can say that this virtual space is rotated, scaled, and zoomed always to give the spectator the best viewpoint. As in a striptease, the space slowly disrobes itself, turning, presenting itself from different sides, teasing, stepping forward and retracting, always leaving something covered so that the spectator must wait for the next shot . . . a seductive dance that begins all over with the next scene. All the spectator has to do is remain immobile.

Film theorists have taken this immobility to be the essential feature of the institution of cinema. Anne Friedberg writes: "As everyone from Baudry (who compares cinematic spectation to the prisoners in Plato's cave) to Musser points out, the cinema relies on the immobility of the spectator, seated in an auditorium."[63] Film theoretician Jean-Louis Baudry, probably more than anyone else, emphasizes immobility as the foundation of cinematic illusion, quoting Plato: "In this underground chamber they have been from childhood, chained by the leg and also by the neck, so that they cannot move and can only see what is in front of them, because the chains will not

61. Quoted in ibid., 215.

62. Ibid., 214.

63. Friedberg, *Window Shopping,* 134. She refers to Jean-Louis Baudry, "The Apparatus: Metapsychological Approaches to the Impression of Reality in the Cinema," in *Narrative, Apparatus, Ideology,* ed. Philip Rosen (New York: Columbia University Press, 1986) and Charles Musser, *The Emergence of Cinema: The American Screen to 1907* (New York: Charles Scribner and Sons, 1990).



let them turn their heads."[64] This immobility and confinement, according to Baudry, enables the prisoners/spectators to mistake representations for their perceptions thereby regressing to childhood when the two were indistinguishable. Rather than a historical accident, the immobility of the spectator, according to Baudry's psychoanalytic explanation, is the essential condition of cinematic pleasure.

Alberti's window, Dürer's perspectival machines, the camera obscura, photography, cinema—in all of these screen-based apparatuses, the subject has to remain immobile. In fact, as Friedberg perceptively points out, the progressive mobilization of the image in modernity was accompanied by the progressive imprisonment of the viewer: "as the 'mobility' of the gaze became more 'virtual'—as techniques were developed to paint (and then to photograph) realistic images, as mobility was implied by changes in lighting (and then cinematography)—the observer became more immobile, passive, ready to receive the constructions of a virtual reality placed in front of his or her unmoving body."[65]

What happens to this tradition with the arrival of a screen-less representational apparatus—VR? On the one hand, VR constitutes a fundamental break with this tradition. It establishes a radically new type of relationship between the body of the viewer and the image. In contrast to cinema, where the mobile camera moves independently of the immobile spectator, now the spectator actually has to move in physical space in order to experience movement in virtual space. It is as though the camera were mounted on the user's head. Thus, to look up in virtual space, one has to look up in physical space; to step forward "virtually" one has to step forward in actuality, and so on.[66] The spectator is no longer chained, immobilized, anesthetized by the apparatus that serves her ready-made images; now she has to work, to speak, in order to see.

At the same time, VR imprisons the body to an unprecedented extent. This can clearly be seen in the earliest VR system designed by Sutherland

64. Quoted in Baudry, "The Apparatus," 303.

65. Friedberg, *Window Shopping,* 28.

66. A typical VR system adds other ways of moving around, for instance, the ability to move forward in a single direction by simply pressing a button on a joystick. To change the direction, however, the user still has to change the position of his/her body.

and his colleagues in the 1960s. According to Howard Rheingold's history of VR, "Sutherland was the first to propose mounting small computer screens in binocular glasses—far from an easy hardware task in the early 1960s—and thus immerse the user's point of view inside the computer graphic world."[67] Rheingold further wrote:

> In order to change the appearance of the computer-generated graphics when the user moves, some kind of gaze-tracking tool is needed. Because the direction of the user's gaze was most economically and accurately measured at that time by means of a mechanical apparatus, and because the HMD [head-mounted display] itself was so heavy, the users of Sutherland's early HMD systems found their head locked into machinery suspended from the ceiling. The user put his or her head into a metal contraption that was known as the "Sword of Damocles" display.[68]

A pair of tubes connected the display to tracks in the ceiling, "thus making the user a captive of the machine in a physical sense."[69] The user was able to turn around and rotate her head in any direction, but could not move away from the machine more than a few steps. Like today's computer mouse, the body was tied to the computer. In fact, the body was reduced to nothing less—and nothing more—than a giant mouse, or more precisely, a giant joystick. Instead of moving a mouse, the user had to turn her own body. Another comparison that comes to mind is the apparatus built in the late nineteenth century by Etienne-Jules Marey to measure the frequency of the wing movements of a bird. The bird was connected to the measuring equipment by wires that were long enough to enable it to flap its wings in midair but not fly anywhere.[70]

The paradox of VR, that it requires the viewer to move in order to see an image and at the same time physically ties her to a machine, is interestingly dramatized in a "cybersex" scene in the movie *Lawnmower Man* (Brett Leonard, 1992). In the scene, the heroes, a man and a woman, are situated in

67. Rheingold, *Virtual Reality,* 104.

68. Ibid., 105.

69. Ibid., 109.

70. Marta Braun, *Picturing Time: The Work of Etienne-Jules Marey (1830–1904)* (Chicago: University of Chicago Press, 1992), 34–35.



the same room, each fastened to a separate circular frame that allows the body to rotate 360 degrees in all directions. During "cybersex" the camera cuts back and forth between virtual space (i.e., what the heroes see and experience) and physical space. In the virtual world represented by psychedelic computer graphics, their bodies melt and morph together, disregarding all the laws of physics, while in the real world each of them simply rotates within his or her own frame.

The paradox reaches its extreme in one of the most long-standing VR projects—the Super Cockpit developed by the U.S. Air Force in the 1980s.[71] Instead of using his eyes to follow the terrain outside the plane and the dozens of instrument panels inside the cockpit, the pilot wears a head-mounted display that presents both kinds of information in a more efficient way. What follows is a description of the system from *Air & Space* magazine:

> When he climbed into his F16C, the young fighter jock of 1998 simply plugged in his helmet and flipped down his visor to activate his Super Cockpit system. The virtual world he saw exactly mimicked the world outside. Salient terrain features were outlined and rendered in three dimensions by the two tiny cathode ray tubes focused at his personal viewing distance. . . . His compass heading was displayed as a large band of numbers on the horizon line, his projected flight path a shimmering highway leading out toward infinity.[72]

If in most screen-based representations (painting, cinema, video) as well as typical VR applications, the physical and virtual worlds have nothing to do with each other, here the virtual world is synchronized precisely with the physical one. The pilot positions himself in the virtual world in order to move through the physical one at a supersonic speed with his representational apparatus securely fastened to his body, more securely than ever before in the history of the screen.

### Representation versus Simulation

In summary, VR continues the screen's tradition of viewer immobility by fastening the body to a machine, while at the same time it creates an

71. Rheingold, *Virtual Reality,* 201–209.

72. Quoted in ibid., 201.

unprecedented new condition by requiring the viewer to move. We may ask whether this new condition is without historical precedent, or whether it fits within an alternative representational tradition that encourages the movement of the viewer.

I began my discussion of the screen by emphasizing that a screen's frame separates two spaces that have *different* scales—the physical and the virtual. Although this condition does not necessarily lead to the immobilization of the spectator, it does discourage any movement on her part: Why move when she can't enter the represented virtual space anyway? This is well dramatized in *Alice in Wonderland* when Alice struggles to become just the right size in order to enter the other world.

The alternative tradition of which VR is a part can be found whenever the scale of a representation is the same as the scale of our human world so that the two spaces are continuous. This is the tradition of simulation rather than that of representation bound to a screen. The simulation tradition aims to blend virtual and physical spaces rather than to separate them. Therefore, the two spaces have the same scale; their boundary is de-emphasized (rather than being marked by a rectangular frame, as in the representation tradition); the spectator is free to move around the physical space.

To analyze further the different logic of the two traditions, we may compare their typical representatives—frescoes and mosaics, on the one hand, and Renaissance painting, on the other. The former create an illusionary space that starts behind the surface of an image. Importantly, frescoes and mosaics (as well as wall paintings) are inseparable from architecture. In other words, they cannot not be moved anywhere. In contrast, the modern painting, which first makes its appearance during the Renaissance, is essentially mobile. Separate from a wall, it can be transported anywhere. (It is tempting to connect this new mobility of representation with the tendency of capitalism to make all signs as mobile as possible.)

But, at the same time, an interesting reversal takes place. Interaction with a fresco or a mosaic, which itself cannot be moved, does not assume immobility on the part of the spectator, while the mobile Renaissance painting does presuppose such immobility. It is as though the imprisonment of the spectator is the price for the new mobility of the image. This reversal is consistent with the different logic of the representation and simulation traditions. The fact that the fresco and mosaic are "hardwired" to their architectural setting allows the artist to create a continuity between



virtual and physical space. In contrast, a painting can be put in an arbitrary setting, and therefore, such continuity can no longer be guaranteed. Responding to this new condition, a painting presents a virtual space that is clearly distinct from the physical space where the painting and spectator are located. At the same time, its imprisons the spectator through a perspective model or other techniques so that she and the painting form one system. Therefore, if in the simulation tradition, the spectator exists in a single coherent space—the physical space and the virtual space that continues it—in the representational tradition, the spectator has a double identity. She simultaneously exists in physical space and in the space of representation. This split of the subject is the tradeoff for the new mobility of the image as well as for the newly available possibility to represent any arbitrary space, rather than having to simulate the physical space where an image is located.

While the representational tradition came to dominate post-Renaissance culture, the simulation tradition did not disappear. In fact, the nineteenth century, with its obsession with naturalism, pushed simulation to the extreme with the wax museum and the dioramas of natural history museums. Another example of the simulation tradition is sculpture on a human scale, for instance, Auguste Rodin's "The Burghers of Calais." We think of such sculptures as part of a post-Renaissance humanism that puts the human at the center of the universe, when in fact they are aliens, black holes uniting our world with another universe, a petrified universe of marble or stone that exists in parallel to our own.

VR continues the tradition of simulation. However, it introduces one important difference. Previously, the simulation depicted a fake space continuous with and extended from the normal space. For instance, a wall painting created a pseudo landscape that appeared to begin at the wall. In VR, either there is no connection between the two spaces (e.g., I am in a physical room while the virtual space is an underwater landscape) or, conversely, the two completely coincide (e.g., the Super Cockpit project). In either case, the actual physical reality is disregarded, dismissed, abandoned.

In this respect, the nineteenth-century panorama can be thought of as a transitional form between classical simulations (wall paintings, human-size sculpture, diorama) and VR. Like VR, the panorama creates a 360-degree space. Viewers are situated in the center of this space, and they are encouraged to move around the central viewing area in order to see different parts of the panorama.[73] But in contrast to wall paintings and mosaics that, after all, act as decorations of a real space, the physical space of action, now this physical space is subordinate to the virtual space. In other words, the central viewing area is conceived as a continuation of fake space, rather than vice versa, as before—and this is why it is usually empty. It is empty so that we can pretend that it continues the battlefield, or the view of Paris, or whatever else the panorama represents.[74] From here we are one step away from VR, where physical space is totally disregarded, and all "real" actions take place in virtual space. The screen disappeared because what was behind it simply took over.

And what about the immobilization of the body in VR that connects it to the screen tradition? Dramatic as it is, this immobilization probably represents the last act in the long history of the body's imprisonment. All around us are the signs of increasing mobility and the miniaturization of communication devices—mobile telephones and electronic organizers, pagers and laptops, phones and watches that offer Web surfing, Gameboys, and similar handheld game units. Eventually, the VR apparatus may be reduced to a chip implanted in the retina and connected by wireless transmission to the Net. From that moment on, we will carry our prisons with us—not in order to blissfully confuse representations and perceptions (as in cinema), but rather always to "be in touch," always connected, always "plugged-in." The retina and the screen will merge.

This futuristic scenario may never become a reality. For now, we clearly live in the society of the screen. Screens are everywhere—the screens of airline agents, data-entry clerks, secretaries, engineers, doctors, and pilots; the screens of ATM machines, supermarket checkouts, automobile dashboards,

73. Here I disagree with Friedberg, who writes, "Phantasmagorias, panoramas, dioramas—devices that concealed their machinery—were dependent on the relative immobility of their spectators" (23).

74. In some nineteenth-century panoramas, the central area was occupied by the simulation of a vehicle consistent with the subject of the panorama, such as a part of a ship. We can say that in this case the virtual space of the simulation completely takes over the physical space; that is, physical space has no identity of its own—not even such minimal negative identity as emptiness. It completely serves the simulation.



and, of course, the screens of computers. Rather than disappearing, the screen threatens to take over our offices and homes. Both computer and television monitors are getting bigger and flatter; eventually, they will become wall-sized. Architects such as Rem Koolhaas design *Blade Runner*-like buildings whose façades have been transformed into giant screens.[75]

Dynamic, real-time, and interactive, a screen is still a screen. Interactivity, simulation, and telepresence: As was the case centuries ago, we are still looking at a flat, rectangular surface, existing in the space of our body and acting as a window into another space. We still have not left the era of the screen.

75. I am referring here to Rem Koolhaas's unrealized project for a new building for ZKM in Karlsruhe, Germany. See Rem Koolhaas and Bruce Mau, *S, M, L, XL* (New York: Monacelli Press, 1995).