perspective in The ...

"The Language of New Media"
by Lev Manovich
M.I.T. Press
© 2001

wh... are ... worl... spati... a gar... geom... ing bi... ders w... of a tre... throug... an illus... world is ... ...ture of the ...because each cylinder is rotating at its own spe... 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.

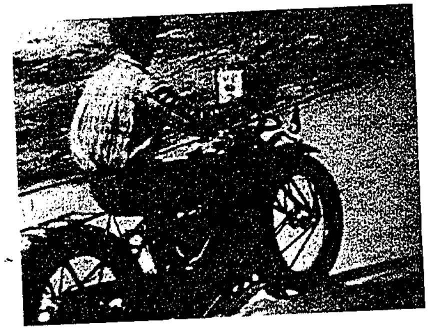
To see that this hybrid language of the cultural interfaces of the 1990s represents only one historical possibility, consider a very different scenario. Potentially, cultural interfaces could completely rely on already existing metaphors and action grammars of a standard HCI, or, at least, rely on them 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 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.

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.

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 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]

---

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).

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 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]

---

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.

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 fixed in 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]

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

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 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, 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.

---

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.

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).