

# Cinema as a Cultural Interface

*author: Lev Manovich*

*year: 1997*

## **The Most Popular Moving Image Sequence of All Times**

Don't you wish that somebody, in 1895, 1897, or at least in 1903, realized the fundamental significance of cinema's emergence and produced a comprehensive record of new medium's emergence? [1] Interviews with the audiences; a systematic account of the narrative strategies, scenography and camera positions as they developed year by year; an analysis of the connections between the emerging language of cinema and different forms of popular entertainment which coexisted with it, would have been invaluable. But, of course, these records do not exist. Instead, we are left with newspaper reports, diaries of cinema's inventors, programs of film showings and other bits and pieces — a set of random and unevenly distributed historical samples.

Today we are living in the midst of an emerging new medium — the metamedium of the digital computer. All information becomes encoded in one code; all cultural objects become computer programs, something which is not only seen, heard or read, but first of all stored and transmitted, compiled and executed. In contrast to a hundred years ago, when cinema was coming into being, we are fully aware of the significance of this new media revolution. And yet I am afraid that future theorists and historians of computer media will be left with not much more than the equivalents of newspaper reviews and random bits of evidence similar to cinema's first decades. They will find that the analytical texts from our era are fully aware of the significance of computer's takeover of culture yet, by and large, they mostly contain speculations about the future rather than a record and a theory of the present. Future researchers will wonder why the theoreticians, who already had plenty of experience analyzing older cultural forms, did not try to describe computer media's semiotic codes, modes of address, and audience reception patterns. If, for instance, they painstakingly reconstructed how cinema emerged out of preceding cultural forms

(panorama, optical toys, peep shows), why didn't they attempt to construct a similar genealogy for the language of computer media at the moment when it was just coming into being, while the elements of previous cultural forms going into its making are still clearly visible, still recognizable before melting into a new unity. Where were the theoreticians at the moment when the icons and the buttons of multimedia interfaces were like a wet paint on a just completed painting, before they became a universal convention and thus slipped into invisibility? Or, at the moment when the designers of *Myst* were debugging their code, converting graphics to 8-bit and massaging QuickTime clips? Or, at the historical moment when a young 20-something programmer at Netscape took the chewing gum out of his mouth, sipped warm Coke out of the can — he was at a computer for 16 hours straight, trying to meet a marketing deadline — and, finally satisfied with its small file size, saved a short animation of stars moving across the night sky, the animation which was to appear in the upper right corner of Netscape Navigator, thus becoming the most widely seen moving image sequence ever — until the next release.

The following is an attempt at both a record and a theory — of the present. Just as film historians traced the development of film language during cinema's first decades, I want to describe and understand the logic driving the development of the language of computer media. It is tempting to extend this parallel a little further and to speculate whether today this new language is already getting closer to acquiring its final and stable form, just as film language acquired its "classical" form during the 1910s. Or are the 1990s more like the 1890s, because future computer media language will be entirely different than the one used today? [2] In either case, by trying to understand which cultural forces are shaping the development of this language, we may be in a better position both to predict its future course as well as to offer different alternatives. For just as avant-garde filmmakers throughout cinema's existence offered alternatives to its particular narrative audio-visual regime, the task of an avant-garde computer artist today is to offer alternatives to the existing language of computer media. This can be better accomplished if we have a theory of how "mainstream" language is currently structured.

Does it make sense to theorize the present when it seems to be changing so fast? It is a gamble. If subsequent developments prove the theoretical projections of this text to be

correct, I win. But, if the language of computer media develops in a different direction than the one suggested by the present analysis, this does not mean that I automatically lose. Rather, the analysis presented here will become a record of possibilities which were heretofore not realized, of the horizon which was visible to us today but later became unimaginable.

We no longer think of the history of cinema as a linear march towards only one possible language, or as a progression towards more and more accurate verisimilitude. Rather, we have come to see its history as a succession of distinct and equally expressive languages, each with its own aesthetic variables, each new language closing off some of the possibilities of the previous one — a cultural logic not dissimilar to Kuhn's analysis of scientific paradigms. [3] Similarly, every stage in the history of computer media offers its own aesthetic opportunities, as well as its own imagination of the future — in short, its own "research paradigm". This paradigm is modified or even abandoned at the next stage. In this paper I want to record the "research paradigm" of new media during its first decade before it slips into invisibility.

## **Cultural Interfaces**

During the 1990s, the cultural role of a digital computer has changed from a tool to a medium. In the beginning of the decade, a computer was still largely thought of as a simulation of a typewriter, a paintbrush or a drafting ruler — in other words, as a tool used to produce cultural content which, once created, will be stored and distributed in its appropriate media: printed page, film, photographic print, electronic recording. By the end of the decade, the computer's public image has begun to shift to one of a universal machine, used not only to author, but also to store, distribute and access all media. All culture, past and present, is beginning to be filtered through a computer, with its particular human-computer interface.

The term human-computer interface (HCI) describes the ways in which the user interacts with a computer. HCI includes physical input and output devices such a monitor, a keyboard, and a 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 this data, i.e., a grammar of meaningful actions which the user can perform on it. An example of this grammar is the commands used in a command-line interface such as DOS and UNIX: copy file, delete file, set date, open port, list directory, and so on.

As the role of a computer is shifting from being a tool to a universal media machine, 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 would like to introduce the term "cultural interfaces" to describe evolving interfaces used by the designers of Web sites, CD-ROM and DVD-ROM titles, multimedia encyclopedias, online museums, computer games and other digital cultural objects.

If you need to remind yourself what a typical cultural interface looked like in 1997, go back in time and click to a random Web page. You are likely to see something which graphically resembles a magazine layout from the same decade. The page is dominated by text: headlines, hyperlinks, blocks of copy. Within this text are few media elements: graphics, photographs, perhaps a QuickTime movie and a VRML scene. The page also includes radio buttons and a pull-down menu which 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 a database trying to match your entry.

For another example of a prototypical cultural interface of the 1990s, you may 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 a book open in the middle, waiting for your mouse click. Next, an element of a familiar Macintosh interface makes an appearance, reminding you that along with being a new movie/book hybrid, *Myst* is also a computer application: you can adjust sound volume and graphics quality by selecting from a usual Macintosh-style menu in the upper top part

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 which 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 it always happens with cultural languages, only a few of these possibilities actually appear viable in a 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 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?

My theory is that there are three key cultural forms which are shaping cultural interfaces in the 1990s. What are these forms? The answer to this puzzle can be found in the opening sequence of *Myst* which activates them before our eyes, one by one. The first form is cinema. The second form is the printed word. The third form is a general-purpose human-computer interface (HCI).

At the time of this writing (1997), it appears that out of the three, the influence of cinema is becoming more and more important. So, despite frequent pronouncements that cinema is dead, it is actually on its own way to becoming a general purpose cultural interface, a set of techniques and tools which can be used to interact with any cultural data. Accordingly, I will devote the largest section of this article to the discussion of the ways in which cinematic techniques structure cultural interfaces.

As it should become clear from the following, I use words "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 words as cultural forms, mechanisms,

languages or media). "Cinema" thus includes mobile camera, representation of space, editing techniques, narrative conventions, activity of a spectator — 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 already found in panoramas, magic lantern slides, theater and other nineteenth-century cultural forms; similarly, since the middle of the twentieth century, they are 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 which have developed over many centuries (some even before the invention of print) and which 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 which follow each sequentially; a table of contents and index.

Modern human-computer interface has a much shorter history than the printed word or cinema — but it is still a history. Its 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 an accepted convention for operating a computer, and a cultural language in their own right.

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

Bringing cinema, the printed word and HCI interface together and treating them as occupying the same conceptual plane has an additional advantage — a theoretical bonus. It is only natural to think of them as belonging to two different kind 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: they offer 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 to record human memory and human experience, mechanisms for 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 may anticipate at first. On the one hand, being a part of our culture now for half a century, HCI already represents a powerful cultural tradition, a cultural language offering its own ways to represent 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 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; the actions include going from page to page linearly, marking individual pages and using table of contents. In the case of cinema, its physical interface is a particular architectural arrangement of a movie theater; its metaphor is a window opening up into a virtual 3-D space.

Today, as media is being "liberated" from its traditional physical storage media — paper, film, stone, glass, magnetic tape — the elements of printed word interface and cinema interface, which previously were hardwired to the content, become "liberated" as well. A digital designer can freely mix pages and virtual cameras, table of contents 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, 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 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 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 are shaping the language of cultural interfaces.

## **I. Printed Word**

In the 1980s, as PC's and word processing software became commonplace, text became the first cultural media to be subjected to digitization in a massive way. But already in the 1960s, two and a half decades before the concept of digital media was born, researchers were thinking about having 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]).

Text is unique among other 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 meta-language of digital 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 a computer uses text as its meta-language, cultural interfaces in their turn inherit the principles of text organization developed by human civilization throughout its existence. One of these 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 is born in the first centuries of the Christian era when the clay tablets and papyrus rolls are replaced by a codex — the 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 a computer. In 1984, Apple introduced a graphical user interface which 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 these pages, as well as to scroll through individual pages. In this way, a traditional page was redefined as a virtual page, a surface which can be much larger than the limited surface of a computer screen. In 1987, Apple shipped popular Hypercard program which extended the page concept in new ways. Now the users were able to include multimedia elements within the 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 more by enabling the creation of distributed documents, where 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. Finally, with HTML, which allows parts of a single page to be located on different computers, the page became even more fluid and unstable.

The conceptual development of the page in digital media can also be read in a different way — not as 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 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 the information is not available all at once, but arrives sequentially, top to bottom, as though the roll is being unrolled.

A good example of how cultural interfaces stretch the definition of a page while mixing together its different historical forms is the Web page designed in 1997 by the British design collective antirom for HotWired RGB Gallery. [7] The designers have created a large surface containing rectangular blocks of texts 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 in a single page.

By the mid-1990s, Web pages included a variety of media types — but they are still essentially pages. Different media elements — graphics, photographs, digital video, sound and 3-D worlds — were embedded within rectangular surfaces containing text. VRML evangelists wanted to overturn this hierarchy by imaging the 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 stretches for thousands of years, I think it is unlikely that it would disappear so quickly.

While the 1990s cultural interfaces have retained the modern page format, they also have come to rely on a new way of organizing and accessing texts which have little precedent within 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, no such relationship of hierarchy is assumed. The two sources connected through hyperlinking have equal weight; they exist on the same level of importance. Thus the acceptance of hyperlinking in the 1980s can be read as a perfect reflection of contemporary culture with its suspicion of all hierarchies, and its aesthetics of collage where radically different sources are brought together within the singular cultural object ("post-modernism").

Traditionally, texts encoded human knowledge and memory, instructed, inspired, and seduced their readers to adopt new ideas, new ways of interpreting the world, new ideologies. In short, the word was always linked to the art of rhetoric. While it is probably possible to invent a new rhetoric of hypermedia, which will use hyperlinking not to distract the reader from the argument (as it is often the case today), but instead to further convince him/her of 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 have classified hundreds of different rhetorical figures. In the middle of

the twentieth century Roman Jakobson, under the influence of computer's binary logic, information theory and cybernetics to which he was exposed at MIT, radically reduced rhetoric to just two figures: metaphor and metonymy. [9] Finally, in the 1990s, the World Wide Web hyperlinking privileged the single figure of metonymy at the expense of all others. [10] The hypertext of the World Wide Web leads the reader from one text to another, ad infinitum. Contrary to the popular image, in which digital media collapses 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 maybe more accurate to think of the resulting object as an infinite flat surface composed from individual texts in no particular order — the antirom design for HotWired. Expanding this comparison further, we can note that Random Access Memory, the concept behind the group's name, also implies the lack of any 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 the 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 orchestrated breakthroughs, cultural interfaces, like RAM itself, bombards the users with all the data at once. [11]

In the 1980s many critics described one of key's effects of "post-modernism" as that of spatialization: privileging space over time, flattening historical time, refusing grand narratives. Digital media, which has 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 novel and cinema with physical movement through space, as witnessed by endless computer animated fly-throughs or computer games such as *Myst* and countless others. In short, time becomes a flat image or a landscape, something to look at or navigate through. If there is a new rhetoric or aesthetic which is 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 through the sand and water, picking up a navigation journal, a rotten fruit, an instrument whose purpose he does

not know; leaving imprints in the sand, which, like computer hyperlinks, follow from one found object to another.

## **II. Cinema**

Printed word tradition which has initially dominated the language of cultural interfaces, is becoming less important, while the part played by cinematic elements is getting progressively stronger. This is consistent with a general trend in modern society towards presenting more and more information in the form of time-based audio-visual moving image sequences, rather than as text. As new generations of both computer users and computer designers are growing 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 next, are being extended to become the basic ways in which computer users access and interact with all cultural data. In this way, the computer fulfills the promise of cinema as a visual Esperanto which pre-occupied many film artists and critics in the 1920s, from Griffith to Vertov. Indeed, millions of computer users communicate with each other through the same computer interface. And, in contrast to cinema where most of its "users" were 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 an email, run basic applications, organize files and so on.

The original Esperanto never became truly popular. But cultural interfaces are widely used and are easily learned. We have an unprecedented situation in the history of cultural languages: something which is designed by a rather small group of people is immediately adopted by millions of computer users. How is it possible that people around the world adopt today something which a 20-something programmer in Northern California has hacked together just the night before? Shall we conclude that we are somehow biologically

"wired" to the interface language, the way we are "wired," according to the original hypothesis of Noam Chomsky, to different natural languages?

The answer is of course no. Users are able to "acquire" new cultural languages, be it 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, it was theater, magic lantern shows and other nineteenth-century forms of public entertainment. Cultural interfaces in their turn draw on older cultural forms such as the printed word and cinema. 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 function. It became an accepted way for interacting with any data which is represented in three dimensions — which, in a computer culture, means literally anything and everything: the results of a physical simulation, an architectural site, design of a new molecule, financial data, the structure of a computer network and so on. As computer culture is gradually spatializing all representations and experiences, they become 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 besides 3-D space. As an example, consider GUI (Graphical User Interface) of the leading computer animation software — PowerAnimator from Alias/Wavefront. [12] 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. In this way, the model of a virtual camera is extended to apply to navigation through any kind of information, not only the one which was spatialized. It is particularly important that the

user is expected to dolly and pan over text as if it was a 3-D scene. Cinematic vision triumphed over the print tradition, with the camera subsuming the page. The Gutenberg galaxy turned out to be just a subset of the Lumières' universe.

Another feature of cinematic perception which persists in cultural interfaces is a rectangular framing of represented reality. [13] Cinema itself inherited this framing from Western painting. Since the Renaissance, the frame acted as a window onto a larger space which was assumed to extend beyond the frame. This space was cut by the frame's rectangle into two parts: "onscreen space," the part which is inside the frame, and the part which is outside. In the famous formulation of Leon-Battista Alberti, the frame acted as a window onto the world. Or, in a more recent formulation of 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." [14]

Just as a rectangular frame of 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 was final, computer interface benefits from a new invention introduced by cinema: the mobility of the frame. As a kino-eye moves around the space revealing its different regions, so can a computer user scroll through a window's contents.

It is not surprising to see that screen-based interactive 3-D environments, such as VRML worlds, also use cinema's rectangular framing since they rely on other elements of cinematic vision, specifically a mobile virtual camera. It may be more surprising to realize that Virtual Reality (VR) interface, often promoted as the most "natural" interface of all, utilizes the same framing. [15] 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. [16] 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 his/her own sight. Yet, it is crucial that in VR one is seeing the virtual world through a rectangular

frame, and that this frame always presents only a part of a larger whole. This frame creates a distinct subjective experience which is much closer to cinematic perception than to unmediated sight.

Interactive virtual worlds, whether accessed through a screen-based or a VR interface, are often discussed as the logical successor to cinema, as 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 the issues of interaction and narrative. So, 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 commonly invoked in new media discussions of the 1990s, indeed represent an extension of cinema or if they rather should be thought of as a continuation of some theatrical traditions, such as improvisational or avant-garde theater. But what undoubtedly can be observed in the 1990s 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 (Virtual Reality Modeling Language). [17]

The creator of a VRML world can define a number of viewpoints which are loaded with the world. [18] These viewpoints automatically appear in a special menu in a VRML browser which 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 which controls how a VRML browser moves from one viewpoint to the next. By default, the virtual camera smoothly travels 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 their own.

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

The area of computer culture where 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 started featuring lavish opening cinematic sequences (called in the game business "cinematics") to set the mood, establish the setting and introduce the narrative. Frequently, the whole game would be structured as an oscillation between interactive fragments requiring user's input and non-interactive cinematic sequences, i.e. "cinematics". [19] As the decade progressed, game designers were creating increasingly complex — and increasingly cinematic — interactive virtual worlds. Regardless of a game's genre — action/adventure, fighting, flight simulator, first-person action, racing or simulation — they 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 sets to create mood and atmosphere. In the beginning of the decade, games used digital video of actors superimposed over 2-D or 3-D backgrounds, but by its end they switched to fully synthetic characters. [20] This switch also made virtual worlds more cinematic, as the characters could be better visually integrated with their environments. [21]

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 -), after a certain event



takes place (car crashes, a fighter being knocked down), it is 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. Finally, Nintendo went even further by dedicating four buttons on their N64 joystick 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 Rider (Eidos, 1996) also use the buttons on the Playstation joystick for changing point of view.

The incorporation of virtual camera controls into the very hardware of a game consoles is truly a historical event. Directing the virtual camera becomes as important as controlling the hero's actions. This is admitted by the game industry itself. For instance, a package for Dungeon Keeper lists four key features of the game, out of which the first two concern control over the camera: "switch your perspective," "rotate your view," "take on your friend," "unveil hidden levels". In games such as this one, cinematic perception functions as the subject in its own right. [22] Here, the computer games are returning to "The New Vision" movement of the 1920s (Moholy-Nagy, Rodchenko, Vertov and others), which foregrounded new mobility of a photo and film camera, and made unconventional points of view the key part of their poetics.

The fact that computer games continue to encode, step by step, the grammar of a kino-eye in software and in hardware is not an accident. This encoding is consistent with the overall trajectory driving the computerization of culture since the 1940s, that being 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. The 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 in 3-D software, and later directly in hardware. [23] As a result, linear perspective became the default mode of vision in digital culture, be it computer animation, computer games, visualization or VRML worlds. Now we are

witnessing the next stage of this process: the translation of cinematic grammar of points of view into software and hardware. As Hollywood cinematography is translated into algorithms and computer chips, its convention becomes the default method of interacting with any data subjected to spatialization, with a narrative, and with other human beings. (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. [24]) Element by element, cinema is being poured into a computer: first one-point linear perspective; next the mobile camera and a rectangular window; next cinematography and editing conventions, and, of course, digital personas also based on acting conventions borrowed from cinema, to be followed by make-up, set design, and, of course, the narrative structures themselves. From one cultural language among others, cinema is becoming the cultural interface, a toolbox for all cultural communication, overtaking the printed word.

But, in one sense, all computer software already has been based on a particular cinematic logic. Consider the key feature shared by all modern human-computer interfaces — overlapping windows. [25] All modern interfaces display information in overlapping and resizable windows arranged in a stack, similar to a pile of papers on a desk. As a result, the computer screen can present the user with practically an unlimited amount of information despite its limited surface.

Overlapping windows of HCI can be understood as a synthesis of two basic techniques of twentieth-century cinema: temporal montage and montage within a shot. In temporal montage, images of different realities follow each other in time, while in montage within the shot, these different realities co-exist within the screen. The first technique defines the cinematic language as we know it; the second is used more rarely. An example of this technique is the dream sequence in *The Life of an American Fireman* by Edward Porter in 1903, in which an image of a dream appears over a man's sleeping head. Other examples include the split screens beginning in 1908 which show the different interlocutors of a telephone conversation; superimpositions of a few images and multiple screens used by the avant-garde filmmakers in the 1920s; and the use of deep focus and a particular

compositional strategy (for instance, a character looking through a window, such as in *Citizen Kane*, *Ivan the Terrible* and *Rear Window*) to juxtapose close and far away scenes. [26]

As testified by its popularity, temporal montage works. However, it is not a very efficient method of communication: the display of each additional piece of information takes time to watch, thus slowing communication. It is not accidental that the European avant-garde of the 1920s was inspired by the engineering ideal of efficiency, experiments with various alternatives, trying to load the screen with as much information at one time as possible. [27] In his 1927 *Napoleon* Abel Gance uses a multiscreen system which shows three images side by side. Two years later, in *A Man with a Movie Camera* (1929) we watch Dziga Vertov speeding up the temporal montage of individual shots, more and more, until he seems to realize: why not simply superimpose them in one frame? Vertov overlaps the shots together, achieving temporal efficiency — but he also pushes the limits of a viewer's cognitive capacities. His superimposed images are hard to read — information becomes noise. Here cinema reaches one of its limits imposed on it by human psychology; from that moment on, cinema retreats, relying on temporal montage or deep focus, and reserving superimpositions for infrequent cross-dissolves.

In window interface, the two opposites — temporal montage and montage within the shot — finally come together. The user is confronted with a montage within the shot — a number of windows present at once, each window opening up into its own reality. This, however, does not lead to the cognitive confusion of Vertov's superimpositions because the windows are opaque rather than transparent, so the user is only dealing with one of them at a time. In the process of working with a computer, the user repeatedly switches from one window to another, i.e., the user herself becomes the editor accomplishing montage between different shots. In this way, window interface synthesizes two different techniques of presenting information within a rectangular screen developed by cinema.

This last example shows once again the extent to which human-computer interfaces — and the cultural interfaces which follow them — are cinematic, inheriting cinema's particular ways of organizing perception, attention and memory. Yet it also demonstrates the

cognitive distance between cinema and the computer age. For the viewers of the 1920s, the temporal replacement of one image by another, as well as superimposition of two images together were an aesthetic and perceptual event, a truly modern and unfamiliar experience. The cut from one image to another was a meaningful, even stressful event, because audiences had to assimilate a sequence in a different fashion than they were previously used to in other cultural forms. [28] Film directors exploited the novelty of this strategy as an effective way of creating meaning. At the end of the century, however, anesthetized first by cinema and then by television channel flipping, we feel at home with a number of overlapping windows on a computer screen. We switch back and forth between different applications, processes, tasks. Not only are we no longer shocked, but in fact, we feel angry when a computer occasionally crashes because we opened too many windows at once.

Cinema, the major cultural form of the twentieth century, has found a new life as the toolbox of a computer user. Cinematic means of perception, of connecting space and time, of representing human memory, thinking, and emotions 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 in a fictional world of a cinematic narrative has become a window in a datascape. In short, what was cinema has become human-computer interface.

I will conclude this section by discussing a few artistic projects which, in different ways, offer alternatives to this trajectory. To summarize it once again, the trajectory involves gradual translation of elements and techniques of cinematic perception and language into a decontextualized 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 contexts of its formation. If in cinema the camera functioned as a material object, co-existing, spatially and temporally, with the world it was showing us, it has now become a set of abstract operations. The art projects described 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 a virtual world's ontology. They also refuse the universalization of cinematic vision by computer culture, which (just as post-modern visual culture in general) treats cinema as a

toolbox, a set of "filters" which can be used to process any input. In contrast, each of these projects employs a unique cinematic strategy which has a specific relation to the particular virtual world it reveals to the user.

In my own project Reality Generator (1996 - ongoing) I directly make points of view a part of the ontology of a virtual world. The world is described as a set of objects and a set of viewpoints attached to different points in space. Some viewpoints are simply XYZ coordinates which do not correspond to anything in particular. Other viewpoints are attached to particular objects: a leaf, a bottle on the ground, a cloud. In this way, every object also becomes the subject, the focalizer of the narrative. [29] Everything can be seen from any position. Modernist techniques of switching between narrators in different parts of the story and re-telling the same events from different points of view are combined with computer's combinatory logic.

In *The Invisible Shape of Things Past* Joachim Sauter and Dirk Lusenbrink of the Berlin-based Art+Com collective created a truly innovative cultural interface for accessing historical data about Berlin's history. [30] The interface de-virtualizes cinema, so to speak, by placing the records of cinematic vision back into their historical and material context. As the user navigates through a 3-D model of Berlin, he or she comes across elongated shapes lying on city streets. These shapes, which the authors call "filmobjects", correspond to documentary footage recorded at the 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 is getting correspondingly thinner.

In following with the already noted general trend of computer culture towards 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 as book pages. The trajectory through time and space were taken by a camera becomes a book to be read, page by page. The records of camera's vision become material objects, sharing the space with the material reality which

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, as it is becoming today in a computer.

Hungarian-born artist Tamás Waliczky openly refuses the default mode of vision imposed by computer software, that of the 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 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 a film in which it is used. In *The Garden*, its subject is the perspective of a small child, for whom the world does not yet have an objective existence. In *The Forest*, the mental trauma of emigration is transformed 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 from his/her environment 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 he moves around, the actual geometry of all the objects around him is transformed, with objects getting bigger as he gets close to him. 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 its perception are fused together.

### III. Human-Computer Interface

The development of human-computer interfaces, until recently, had little to do with cultural applications. Following some of the main applications from the 1940s until the early 1980s, when the current generation of GUI (Graphic User Interface) was developed and reached the mass market together with the rise of a PC (personal computer), we can list the most significant: real-time control of weapons and weapon systems; scientific simulation; computer-aided design; finally, office work with a secretary as a prototypical computer user, filing documents in a folder, emptying a trash can, creating and editing documents ("word processing"). Today, as the computer is starting 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. Thus, 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 may try to recreate "the museum experience" by presenting a navigable 3-D rendering of a museum space, while still resorting to hierarchical menus to 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" which is not directly traceable to either old cultural forms or general-purpose HCI, the designers are still relying on HCI convention in one case - the use of a pull-down menu to switch between different maps of Berlin.

In general, cultural interfaces of the 1990s try to walk an uneasy path between the richness of control provided in general-purpose HCI and an "immersive" experience of traditional cultural objects such as books and movies. Modern general-purpose HCI, be it MAC OS, Windows or Unix, allow their users to perform complex and detailed actions on the digital data: get 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 the imaginary universe whose structure is fixed by the author. Cultural interfaces attempt to mediate between these two fundamentally different and ultimately non-compatible 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 of 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 their 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 links. 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 to act upon; and the more recent convention of graphical human-computer interfaces which, by dividing the computer screen into a set of controls with clearly delineated functions, essentially treats 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 an illusionary space and image as an instrument for action. [31]

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 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 he/she encounters an icon which looks like a magnifying glass it can be used to zoom on documents. In contrast, modern culture (including its "post-modern" 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 contextualized. 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 is striving 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 an 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 artistic 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 a car, plane or any other complex machine. [32] Finally, on yet another level, the traditions of the printed word and of cinema also compete between themselves. One pulls the computer screen towards being dense and flat information surface, while another wants it to 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 don't 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 which uses it quite successfully: 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 don't have to learn yet another cultural interface, search for hyperlinks hidden in images or navigate through a 3-D environment. Reading these files required simply opening them in standard Macintosh SimpleText, one by one. The effect of this simple technique is remarkable. 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 analyzed here illustrate, cultural interfaces try to create their own language rather than simply using 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 a weapon or to analyze statistical data, and it is another to use it to represent cultural memories, values and experiences. The interfaces developed for a computer in its functions of a calculator, control mechanism or a communication device are not necessarily suitable for a computer playing the role of a 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 a computer: its flexibility in displaying and manipulating data, interactive control by the user, and the 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 don't 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 which underwent little changes 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 can keep on changing forever. But there is one thing we can be sure of. We are witnessing the emergence of a new cultural code, something which will be at least as significant as the printed word and cinema before it. We must try to understand its logic while we are in the midst of its natal stage.

## **References:**

[1] I am very grateful to Laura Nix for her help with editing this paper and many valuable suggestions.

[2] For an analysis of the parallels between the language of the nineteenth-century moving image presentations and the language of computer multimedia during the first half of the 1990s, see my "What is Digital Cinema?", in *The Digital Dialectics*, edited by Peter Lunenfeld (Cambridge, Mass.: The MIT Press, 1988).

[3] Thomas S. Kuhn, *The Structure of Scientific Revolutions* (2nd. ed. Chicago: University of Chicago Press, 1970).

[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, Pennsylvania: Carnegie Mellon University, Human-Computer Interaction Institute, 1996).

[5] <http://www.xanadu.net/the.project>, accessed December 1, 1997.

[6] XML which is supposed to replace HTML on the World Wide Web will enable any user to create his/her customized markup language. Thus, the next stage in digital media culture will involve authoring not simply new documents but new languages. For more information on XML, see <http://www.ucc.ie/xml>., accessed December 1, 1997.

[7] <http://www.hotwired.com/rgb/antirom/index2.html>, accessed December 1, 1997.

[8] See, for instance, Mark Pesce, "Ontos, Eros, Noos, Logos," keynote address for International Symposium on Electronic Arts 1995, <http://www.xs4all.nl/~mpesce/iseakey.html>, accessed December 1, 1997.

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

[10] XLM promises to diversify types of links available to include bi-directional links, multi-way links and links to a span of text rather than a simple point. See <http://www.ucc.ie/xml>.

[11] 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 being presented.

[12] See

[http://www.aw.sgi.com/pages/home/pages/products/pages/poweranimator\\_film\\_sgi/indx.html](http://www.aw.sgi.com/pages/home/pages/products/pages/poweranimator_film_sgi/indx.html), accessed December 1, 1997.

[13] In *The Address of the Eye* Vivian Sobchack discusses the three metaphors of frame, window and mirror which underlie modern film theory. The metaphor of a frame comes from modern painting and is central to formalist theory which is concerned with signification;. The metaphor of 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 a 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. Vivian Sobchack, *The Address of the Eye: a Phenomenology of Film Experience* (Princeton: Princeton University Press, 1992).

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

[15] 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*, edited by Woodrow Barfield and Thomas Furness III (New York and Oxford: Oxford University Press, 1995), 175-257.

[16] See Kocian and Task for details on 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 30-50[o].

[17] The following examples refer to a particular VRML browser — WebSpace Navigator 1.1 from Silicon Graphics, Inc. Other browsers have similar features.

<http://webpace.sgi.com/WebSpace/Help/1.1/index.html>, accessed December 1, 1997.

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

[19] For a more detailed analysis of this narrative structure, see my article, "The Aesthetics of Virtual Worlds: Report from Los Angeles," in CTHEORY [www.ctheory.com](http://www.ctheory.com).

[20] Examples of an 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).

[21] Critical literature on computer games, and in particular on their language, remains very slim. Useful facts on history of computer games, descriptions of different genres and the interviews with the 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 and Company, 1997).

[22] *Dungeon Keeper*, MS-DOS/Windows 95 CD-ROM (Bullfrog Productions, 1997).

[23] For a more detailed discussion of the history of computer imaging as gradual automation, see my articles "Mapping Space: Perspective, Radar and Computer Graphics," in *SIGGRAPH '93 Visual Proceedings*, edited by Thomas Linehan, 143-147 (New York: ACM, 1993); and "Automation of Sight from Photography to Computer Vision," in *Electronic Culture: Technology and Visual Representation*, edited by Timothy Druckery and Michael Sand (New York: Aperture, 1996).

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

[25] Overlapping windows were first proposed by Alan Kay in 1969.

[26] The examples of *Citizen Kane* and *Ivan the Terrible* are from Aumont et al., *Aesthetics of Film*, 41.

[27] On the ideal of engineering efficiency in relation to the avant-garde and digital media, see my article "The Engineering of Vision and the Aesthetics of Computer Art," *Computer Graphics* 28, no. 4 (November 1984): 259-263.

[28] The same novelty made possible surrealism.

[29] On the concept of focalization in narrative theory, see Mieke Bal, *Narratology: Introduction to the Theory of Narrative* (Toronto: University of Toronto Press, 1985).

[30] See [http://www.artcom.de/projects/invisible\\_shape/welcome.en](http://www.artcom.de/projects/invisible_shape/welcome.en), accessed December 1, 1997.

[31] 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, as analyzed by Svetlana Alpers in her *The Art of Describing*. In the chapter entitled "Mapping Impulse" she discusses how a Dutch painting of this period functioned as a combined map/picture, combining different kinds of information and knowledge of the world. See Svetlana Alpers, *The Art of Describing: Dutch Art in the Seventeenth Century* (Chicago: University of Chicago Press, 1983).

[32] This historical connection is illustrated by popular flight simulator games where the computer screen is used to simulate the control panel of a plane, i.e. the very type of object from which computer interfaces have developed. The conceptual origin of 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.