SIMULACRUM: Situated Memory for Architectural Space

Parul Shailesh Vora
BA Cognitive Science
minor in Computer Science and Visual Studies
University of California at Berkeley, 2001

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of
Master of Science in Media Arts and Sciences at the
Massachusetts Institute of Technology
September 2003

Author: Parul Shailesh Vora
Program in Media Arts and Sciences
August 8th, 2003

Certified by: V. Michael Bove, Jr.
Principal Research Scientist, MIT Media Laboratory
Advisor

Accepted by: Andrew Lippman
Chair, Departmental Committee on Graduate Students
Program in Media Arts and Sciences

Simulacrum: Situated Memory for Architectural Space

Parul Shailesh Vora

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning on August 8th, 2003
in partial fulfillment of the requirements for the degree of
Master of Science in Media Arts and Sciences

ABSTRACT

The immediacy and transportability of digital images has changed the spatial and temporal relationship between the viewer and the image. The malleability and large volume of these images affords us the ability to set up new such relationships. This thesis introduces a system that creates an asynchronous channel of connection and interaction by allowing two people or two groups of people to simultaneously inhabit a temporally neutral space. Construed as an elastic collective memory, the system intelligently documents audio and visual activities in a social space. This data is dynamically recomposed and manifested in the present as an interactive display environment that composites the past with the present, collapsing the temporal gap between them.

Thesis Advisor: V. Michael Bove, Jr.
Title: Principal Research Scientist, MIT Media Laboratory
This work was supported by the Digital Life Consortium

SIMULACRUM: Situated Memory for Architectural Space

Parul Shailesh Vora

V. Michael Bove, Jr.
Principal Research Scientist of Media Arts and Sciences
MIT Media Laboratory - Object-Based Media
Thesis Advisor

William J. Mitchell
Professor of Architecture and Media Arts and Sciences
MIT Media Laboratory - Smart Cities
Thesis Reader

Joseph A. Paradiso
Associate Professor of Media Arts and Sciences
MIT Media Laboratory - Responsive Environments
Thesis Reader

ACKNOWLEDGEMENTS

Mike Bove, my Advisor, for giving me the opportunity to come to the Media Lab and for giving me the opportunity and help to make the best of it every day after that I've spent here.

Mummytron, Daddytron, and Deep-Sea-Diver, for being my fortune, stability, and light. And for every other reason you could possibly imagine--even including being my beauty treatments, condensed milk transporter and "easy/hard" playmate, my secret Vernors supply and dying plant, and my blue moon and giggles.

Isis Guru (you are my friend), for being the kind of person that e-mails me answers to my questions from 5,000ft. cruising at 500mph.

Scooby, Bahktiar, Tim, Dan, Jeremy, Ali, The Tower of Power, for holding my hand through my introduction to electronics.

Bill Mitchell and Joe Paradiso, my Readers, for reading and everything that goes with that.

The Digital Life Consortium, for making my education something to add to that "Get Free Stuff" book.

Jim, "the hammer," for smoking vegetables, putting the FUN back in FUNdamentalist, reminding me to reinstall, and for defying the paperclip by being a constant source of information and help.

Sammy Davis, Jim, Surj, Darren, Willy, Jacky, Yi, and Ken for making Object Based Media something exciting to be a part of. Especially to Sammy Davis and Jim for making it feel like a team.

Daved, for having an 's' in your name, like a philosopher, making amazing telepathic gyroscopes, introducing me to your friends Borges and Lem way back when, for falling into a pool of nitrogen and dying, and for knowing that you can't smear gravy on communism.

Dr. Bot, for knowing how to true wheels, reinventing my concept of the jetpack, being able to form metal to your will, teaching me the Japanese word for adventure, and for your spelunking enthusiasm in both sparkly and non-sparkly terrain.

K Town Ninja, for poisoning my tea, knowing how to put a headphone splitter to good use, being no one, for the red cross, and for telling me about "those things that people do." (I can still see it).

NickYu, for being a rockstar and for being dumb and ugly. Oh yeah, and for the guitar, the best evil master genius prize.

ACKNOWLEDGEMENTS

M, for psychoanalyzing pinhead, knowing when not to bother gravity, and for wearing tinfoil hats.

Crock and Clement, for sharing my love of gold blocks, being my dungeon master, and being too hot to handle (and cutting that mullet).

G, and sometimes Chad, for knowing how to do things with retinas and G-proteins, thinking that anything can be improved with a pulley system, and bringing me Cheeseboard Pizza.

Bipolar Bear, Neek, for being from the Arctic Circle, not knowing what a knock-knock joke is, looking up molasses in the dictionary, always ordering the dried cranberries, and for sharing Murzik and Muzak.

Shwee and Megan, in Bags and on Bikes, for being XX-rated for twice the adventure and fun, for being able to flap your lip at high frequencies and for bringing down the house with your lightning-speed fixed gear.

aaronedsinger, for always busing your table, putting a servo-motor to good use, and for letting me use this photo "one last time."

Ben and RyanTxu, for giving me Serre and Lapnap, and other things, but more for Serre and Lapnap.

Cian and Jonah, for keeping me connected.

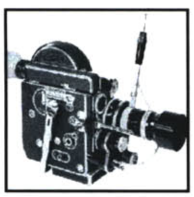
Pat, Linda, John DiFrancesco, Tom, and Brian, for keeping the Media Lab running which kept me running.

Gardeners, for being a group of people that it was nice to see more often than others. And for pretending that you couldn't hear everything that I was saying all the time. Because I'm loud.

And some people just for being friends, Aisling, Amanda, Amy, AndrewF, AndrewS, BenFry, Brant, Cameron, Carson, Cat, Christine, Csik, DanC, DanP, Diana, Dimitris, Eric, Gian, JamesP, Jeana, Justin, Lily, Maeda, Michael, Michelle, Monzy, Ramesh, Saoirse, SimonG, Stacie, Tad, Brian Alex Miller, Maximillian, Steven, Jessica Howe, and BrianTQuinn.

Not to mention the fried dough man, Mr. Klein, my stereo, Lorem Ipsum, any one that's ever built a human powered air travel machine or used a tin can phone, the inventor of shrinky dinks, John Lurie, everyone in an 80s sitcom, Willy Wonka, swingsets, people that know how to break dance, Edwin Land and stores that sell the Holgaroid, cilantro farmers, 15th century explorers, people that like to wear costumes, and anyone that I've ever played a practial joke on or told is smelly or eats dirt.

TABLE OF CONTENTS

I N T R O D U C T I O N

INTRODUCTION

+ Motivations

People are compelled to document their lives and the lives of others. The artifacts we create with writing tools, musical instruments, and cameras serve both as a record of the events and as interface and stimulation for remembering and reliving them. These encapsulated audio-visual experiences are generally also subjected to archival and organizational methods, giving rise to published periodicals, photo albums, and more recently, online blogs, and stashes of Compact Flash™ cards and .avi files. The organizational methods are employed to facilitate personal access to these materials and distribution of them to family, friends, the public, and even strangers. This distribution enhances the sense of connectedness between people that is a seemingly necessary experience of our physical and social world. It is a form of communication not unlike language and gesture, that instinctually eases not only social awareness, but also the benefits of social awareness (i.e. productivity, collaboration).

++ The Electronic Image

The devices for image capture have long since switched from mechanical to chemical to computational. Developments in technology have, in turn, developed the language that they express. What began as capturing a moment in time with specific coordinates in space progressed into the cinematic image. Where photography manipulated space with its depth of field, the cinema explored moments and intervals with its "depth of time." Like most technological developments, it both reflects and transforms social phenomena and constructs--in this case space, time, perception, movement, expression, and subjectivity—and is part of the feedback loop affecting the culture that creates them. The latest development is the "electronic image"—instantly accessible and infinitely reproducible—which is finding its place in the vocabulary of our visual literacy.

Figure 1a-c. Some photographic processes. 1a) Drawing done by the author with a camera lucida. 1b) Darkroom enlarger. 1c) Image from a surveillance camera.

The electronic image is part of the trend that (audio and) visual information is more often than not created, exchanged, distributed, retrieved, and re-used by computational systems. What began as television, which sought to convey and create experiences instead of documenting them, led to the digital cameras and the internet which provided unlimited democratic connectivity and visibility which has resulted in a multi-media overexposure that has manifested itself as reality television shows, 24 hour CNN coverage showing us live views of Tel Aviv as seen from a missile, web and surveillance cameras, and technology mediated communications. The industrialization of vision has led to its democratization.

In her essay, "Delegated Perception," Monica Jaeckal makes the argument with the help of Walter Benjamin that what could be photographed or filmed must have been in front of the camera lens, which set up a relationship between the present viewer and the past moments of space or time which were represented. The electronic image has no such stamp of authenticity if the mode of creation is not evident, owing to its ability to be created, altered, manipulated, and represented in any number of ways (Mitchell, 1992). The vagueness offers both a liberty for expression—lending itself to flexible structures and non-linear timelines--and a problem for representation—offering no clear point of view. We no longer think of these technologies (computer vision, radar, infrared imaging) as showing us what we cannot see, we think of them *seeing* that which we cannot see. What started off as tools for visualization, have developed after the introduction of the computer into tools for vision.

++ Automated Perception

Paul Virilio, in his Vision Machine, argued that "sightless vision," or automated perception, is the latest inevitable stage in the history of the "logistics of perception," his term for the operational agenda according to which perception is appropriated, delimited, and further produced by means of various technologies [controlled by those in political power] (Virilio, 1994). The process has begun, as Lev Manovich points out in his "Metadating the Image." (Manovich, 2002) Metadata is the data about the data (i.e. a keyword or value assigned to an image in a media database), which allows computers to "see" and retrieve data, move it from place to place, compress it and expand it, connect data with other data, and so on. Its creation was the initial scientific response to the dramatic increase of media data being produced and made available. But as Manovich suggests, this has created a modern struggle between the visual data itself (the images) and their creators and masters (humans):

> "The later want to control images: make new images which would precisely communicate the intended

> meanings and effects; yield the exact meanings contained in all the images already created by human cultures; and, more recently, automate these and all over possible image operations by using computers. The former can be said to "resist" all these attempts."

The struggle, in my mind, is two fold because the type of control that humans seek is two fold – one individual, which Manovich highlights, and the other societal. Computers have made image capture, storage, manipulation, and transmission easier and more accessible than ever before, but has not been appropriated yet to deal with the side effects of the burst of production it spurred. Semiotic and hermeneutic issues aside, there is no structure or language within the computer to help in describing and accessing the large quantities of images being generated by digital scanners, cameras, and libraries, not to mention web and surveillance cameras. Manovich is primarily concerned with metadata and database structure as a solution to untap the creative potential these large quantities of media content present. It is this idea that gave rise to this work. But he neglects to address is the potential for this technology to be used for a political agenda as a solution to control the images of surveillant images to govern society that Virilio saw, which needs to be considered in conduction of such work.

As Kuhn explains in his "Structure of Scientific Revolution," there is a genetic aspect between the parallel of scientific and political development (Kuhn, 1996). The majority of cameras in public spaces are installed as safety devices and are thus considered public surveillant devices. From personal observation alone with the abundance of naked, unmanned cameras scattered around the Media Lab, I can see that in many, but not all, they inspire a panoptic paranoia. In "Foucault and Databases: Participatory Surveillance," Mark Poster analyzes how databases, as a means of communicating information digitally become Superpanopticons that allows the self to be surveilled without the "real" self ever noticing. With the introduction of computer vision techniques and associated metadata to these cameras—this dataveillence-- not only does the panopticon surveil in the way that it physically watches the person, but also in the documents that the institution keeps on the person being watched. It does not have walls, windows, towers or guards, yet by the way citizens participate in the social milieu they get inserted in to the database and become part of the social gaze. Poster sees the digitalization

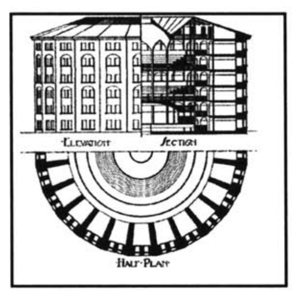


Figure 2a-b.
Are intelligent databases of surveillant images a new kind of panopticon? 2a) Library card catalog system. 2b) Jeremy Bentham's plan for the panoptic prison.

of information as making it easier for people with capitalist interests to distribute certain information to people. The techno-optimistic view of this digitalization is to see it as easing the democratic spread of information, given that the public has, wants, or knows about their access to this information.

++ Augmented Space

Figure 3.
Diller + Scofidio's *Jumpcuts* for a San Jose cinema, where 12 LCDs were suspended on the exterior, displaying alternating live, mediated images mixed and un-mediated images, acting as a theater curtain between the spectacle of the real world and of that of the cinema.

The problem is, the public does not have control over "public" cameras. The solution being simple, give them access. The parallel development of "surveillant devices" (small and cheap cameras), what Lev Manovich refers to as cellspace technologies (sensors, GPS devices, mobile communication tools), and computer/video displays, is transforming our physical space into dataspace—an overlaying layers of data over the physical space commonly referred to as "augmented space." In his text *Iconography and Electronics Upon a Generic Architecture*, Robert Venturi proposes that architecture should look to returning to its traditional definition as communicative surfaces to once more make architecture relevant to those who use it. "Of course, if the messages communicated by traditional architecture were static and reflected the dominant ideology, today electronic dynamic interactive displays make possible for these messages to change continuously and to be the space of contestation and dialog, thus functioning as the material manifestation of the often invisible public sphere." (Manovich, 2002).

This seems to me an important direction for the uses of new such technologies. Our communications and interactions with each other are increasingly technology-mediated, whether through email or video conferencing, and physical proximity and shared meeting spaces are loosing their significance in personal encounters. By giving space the technology-aided ability to collect, store, retrieve and display visual and audio information (with the addition of cameras, sensors, and display devices), it embeds in its architecture that information as part of its structure and history. The augmented space becomes an evolving artifact of the community and a member of the community itself, owing to its ability to store this information created by both individuals and the community and using this information to interact with them—it is the body for the otherwise incorporeal eye.

+ Approach

In 1967, Guy DeBord boldly stated that "In societies where modern conditions of production prevail, all of life presents itself as an immense accumulation of spectacles. Everything that was directly lived has moved away into a representation." (DeBord, 1967) Surveillance has become the new spectacle, noted most notably by Michel Foucault and DeBord himself in 1988. DeBord sees the core of the spectacle as the annihilation of historical knowledge—in particular the destruction of the recent past. As Jonathan Crary writes, "in its place there is the reign of a perpetual present."

++ Perception and Memory

I find myself unable to distinguish the cognitive process of perception from that of memory. Even with my Cognitive Science, Photography, and Computer Science background, I find Henri Bergon's the most descriptive for the experience of what he calls "pure" memory. Bergson cites the image as the currency of exchange between perception and memory and treats it as a mode of transportation between the present (or the perceiving) and the past (or the remembering). "Memory thus creates anew present perception," he says. Deleuze later sought to describe a similar such moment in cinema in which memory has the capacity to rebuild the object of perception. According to his analysis, "the image is not an icon or simulacrum representing something existing in the world; it is rather the perceptual correlative of actions in and reaction to a milieu" as Bergson prescribed. This milieu, John Johnston says, is now defined by a variety of agents and sub-agents in human machine systems, and this author believes that temporal code is disappearing with the increasing immediacy of the electronic image.

One of the schools in Tlön has reached the point of denying time. It reasons that the present is undefined, that the future has no other reality than as present hope, that past is no more than present memory. Another maintains that the universe is comparable to those code systems in which not all the symbols have meaning, and in which only that which happens every three hundredth night is true.

-- Jorge Borges, "Tlön, Uqbar, Orbis Tertius"

New trends and developments in media technologies, particularly in the compositing of images and the capability to record images continuously, afford us the ability to re-establish this code. "If both traditional arts and modern media are based on sampling reality, that is, representing/recording only small fragments of human experience, digital recording and storage technologies greatly expand how much can be represented/recorded. This applies to granularity of time, the granularity of visual experience, and also what can be called "social granularity" (i.e., representation of one's relationships with other human beings.)....The social universe no longer needs to be sampled but can be modeled as one continuum." (Manovich, 2002)

++ Active Capture and Social Memory

But perhaps a complete recreation of memory is not necessary for its representation—that is memory is not such a rigid archive, and the continuum can afford greater elasticity. True, media capture devices and their memory capabilities are undergoing daily re-invention, but as Marc Davis points out, the apparatus is changing, but "the interaction paradigms for media capture have not." Mike Bove's Object-Based-Media Group at the MIT Media Lab is focused on rethinking and reinventing the media capture, production, and automation process in development of electronic visual communication. Parallel efforts such as the Garage Cinema Research Group at UC-Berkeley are agreeing on the need for media metadata, active capture, and adaptive (or alternatively, personalized) environments in to help in the filtration and presentation of media applications with various purposes.

* *Media Metadata*: a means to describe media content and structure, designed to facilitate media retrieval, recombination, and reuse.
* *Active Capture*: combining capture, (image) processing, and interaction to enable the communication and interaction among the capture device, human agents, and the environment shared by the device and agents; accomplished by integrating technology and techniques from computer vision, human-computer interaction design, and media production.
* *Adaptive Environments*: using annotated assets and automated media editing functions for the mass customization and personalization without any effort by the user.

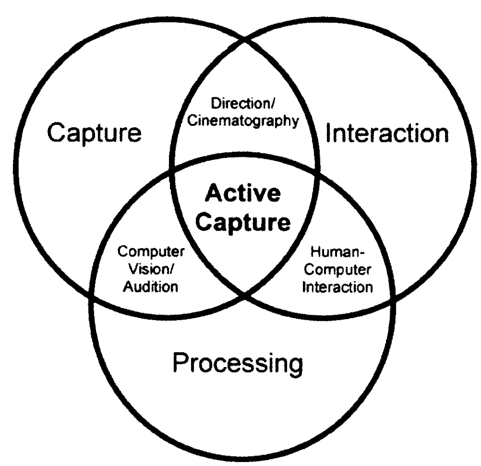


Figure 4.
Active Capture Model by Marc Davis and the Garage Cinema Group at Berkeley.

The active capture model is subsumed by current research on "smart cameras," which aims to create a single device with on board sensing, processing and memory to carry out the video analysis for the detection of objects, human presence, gestures, and human activity, among other things. "The majority of the research in human identification concentrates on algorithm development and is done in non-real time." (Ozer, Wolf, 2001). For environmental (video) applications that appropriate the detection and recognition of objects and people for the purposes of interacting with them, real-time video analysis is required.

Although such smart cameras are not quite product-ready, in concept they can be applied to the selective capture and adaptive representation of individual and social memory. That is, the information they provide could be used to simultaneously record particular activities and determine the manner in which these events are reconstructed dependent on the current state of activity. Maurice Halbwachs argues that a variety of factors that derive from the social arena people inhabit are as important as the individual in the function of memory. He introduced the term 'mémoire collective' or collective memory and stressed how strongly social states influence not only people's personal memories of their own lifetimes, but also a community's shared memories of the past (Kohli-Kunz, 1973). Going beyond Halbwachs's argument, cognitive psychologists, neurobiologists, and sociologists have recently proposed that human memory works radically differently from the traditional archive model and is in fact constructed in the human brain according to the conditions and viability of the remembering subject (Bolles, 1998). Memory of the past, it is argued, is not only influenced but also constituted by physical and social contexts of the present, necessitating the adaptive environments model. As a consequence whether or not a particular event or process remembered corresponds to the actual past becomes less important than the specific conditions under which such memory is created and constructed as well as the personal and social implications of memories held.

+ Thesis Overview

The remainder of this thesis is comprised of four chapters: Background, Preliminary Work, Implementation, Evaluation & Future Directions, and Conclusions.

The chapter named Background gives a historical foundation for this document, an overview of the research areas that are pertinent to this work, including Media Spaces, Responsive Environments, Mediated Architecture, and Video Databasing. Examples of research in the individual areas as well as other research that encompasses these areas are given both as stories of success and with shortcomings that helped shaped this work from the onset. The aim of establishing a histo-theoretical basis for this work to provide a comparative evaluation of concepts and technologies that relate to it.

The document is not the fortunate tool of a history that is primarily and fundamentally memory. In our time, history is that which transforms documents into monuments. In that area where, in the past, history deciphered the traces left by men, it now deploys a mass of elements that have to be grouped, made relevant, placed in relation to one another to form totalities; it might be said, to play on words a little, that in our time history aspires to the condition of archaeology, to the intrinsic description of the monument.

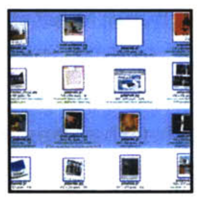
-- Michel Foucault, "Archaelogy of Knowledge."

Included next is the chapter entitled Preliminary Work which highlights projects that I have helped with or been a part of at the Media Lab, including AudioPrint, Reflexion, DoubleExposure, MotoFoto, and Poladroid. It is only in writing this document that I can see that some of the same ideas keep surfacing in my work, at each iteration dropping some facets and gaining others.

The Implementation chapter follows with a technical step by step of software, hardware and physical design stratagem employed in creating Simulacrum to give illustrations of mistakes made and fruitful developments that can be used to direct future research in the area and highlight areas of possible potential, which are discussed at length in the fifth chapter, Evaluation.

The thesis is concluded by outlining the reasons for the limitations and successes reached in view of the project goals that were put forth, all of which are given as reasons of how this thesis is a contribution to an interested community.



B A K C G R O U N D

BACKGROUND

I have an affinity for naming things that goes beyond pets and cars and computers; my bike, my plants, my camera all are ascribed a name. The name for this project-Simulacrum- was intended as a reexamination of Baudrillard's text Simulacra and Simulation. Baudrillard uses the concepts of the simulacra—a copy without an original—to understand the development of the postmodern "to the extent [the simulacrum] addresses the concept of mass reproduction and reproducibility that characterizes our electronic media culture."

Baudrillard begins with a metaphor drawn from a Borges fable in which a cartographer draws a map in such detail that it ends up extending across the entire actual territory he was trying to represent. This metaphor, Baudrillard argues, can no longer hold with the current (1988) reality because we are existing in a state of "hyperreality," in which abstractions are no longer "the map, the double, the mirror, or the concept." The breakdown began with the proliferation and mass production and reproduction of copies that has inundated us in a surplus of images. Tendencies in our culture to surround ourselves with these endless reproductions of the real have annihilated the distinction between reality and its representation, what has been produced and what produces itself. These distinctions have been so thoroughly blurred, that there is no longer any way of differentiating between a true "experience" and its mediated version.

Two decades later, there is a consensus that Baudrillard is overly pessimistic and is fueled by a nostalgia for the old reality, the original simulacra. Deleuze and Guttari have since come along to show us that the simulacrum is not a form of deception[1], but rather a beginning masking the advent of a whole new vital dimension (Deleuze, 1983). What Baudrillard does not account for is the dependence of each "phase" on the previous one--the pasts role in the present—fueled by the sense that each image of reality can in turn affect reality. His model presupposes his conclusion. It is believed by this author that in building on and representing prior realities we are not destroying the relation between real and representation, but rather absorbing it to build a richer such relation. Things are not replaced, merely enveloped. I read Baudrillard's procession not as a pessimistic road map to the apocalypse but as an optimistic empowerment for techno-scientific and artistic development.

1 Plato distinguishes between good copies and bad copies. Good copies are based upon the degree to which the representation of appearances resemble ideal forms or Ideas. Bad copies, on the other hand, are imitations of appearances that while seeming to perfectly mimic reality are upon close inspection not even like the originals they profess to resemble. For Plato it is bad copies that give rise to the simulacra; to a false representation that challenges the primacy of sameness linking appearances to models to Ideal Forms. And it is the simulacra that Plato represses in the search for a knowledge and truth that enlightens rather than deceives. In turn, argues Deleuze, Plato's decision to exorcise the simulacra from the order of representation constructs a legacy in Western culture of repressing difference in favour of sameness, to infuse the copy with the power to affect the original. Thus to assert the primacy of the simulacra is not to give into a world of degraded copies, but in Deleuze's words to "render the order of participation, the fixity of distribution, the determination of hierarchy impossible."

This sentiment is expressed by Jim Hollan and Scott Stornetta in their article "Beyond Being There," in reference to using technology to reproduce face-to-face communications. The article is a critique of the general trend in the telecommunications field to focus their attentions on recreating the audio and video channels as a simulation of physical proximity. They state that "any system which attempts to bring those that are physically distant into a physically proximate community by imitating physical proximity will always keep the former at a disadvantage. This is not because of the quality of the systems, but because of what they attempt to achieve." Just as it is believed that the point is not for the simulacrum to imitate and become the thing it simulates, but to temporarily use this mask for other goals and its own proliferation, Hollan and Stornetta argue that we must look to make communication tools that go "beyond being there." By this they mean that focus should be shifted toward the development of tools that offer a mode of interaction that is an enhancement of the original model—face to face communication. "Real progress on the telecommunication problem" necessitates the development of "tools that people prefer to use even when they have the option of interacting as they have heretofore in physical proximity."

I found myself exploring this idea of advancement again in reading Foucault's The Order of Things. Foucault's method of discourse on historical development focuses on discontinuities and interruptions in the history of thought, which he says

> "suspend the continuous accumulation of knowledge, interrupt its slow development, and force it to enter a new time." They direct historical analysis away from the search for silent beginnings, and the never-ending tracing back to the original precursors, towards the search for a new type of rationality and its various affects. "They [breaches in the continuity of thought] show that the history of a concept is not wholly and entirely that of its progressive refinement, its continuously increasing rationality its abstraction gradient, but that of its various fields of constitution and validity, that of its successive rules of use, that of the many theoretical contexts in which it was developed and matured."

In reading this, I found a method with which to research, examine, and even appropriate prior research that could be absorbed into my project goals in hand. It is only fitting that this method is analogous to the type of functionality that I was aspiring to give my system to progressively refine and mature our social memory. Such a method applied to Baudrillard's model would no longer imply that science destroys what it studies, but rather that this process is a necessary factor of its development—its progress. There are a number of existing and current research projects in media spaces and mediated architecture, responsive environments, video databasing, and architectural interfaces that have become part of the development of this thesis project, not as something to refine or to even build off of, but as something that was contained by the motivating idea and outlined research goals.

+ Media Spaces and Responsive Environments

The term "Media Space" was coined by R. Stults and his colleagues at Xerox Parc in 1986 as being "An electronic setting in which groups of people can work together, even when they are not resident in the same place or present at the same time. In a media space, people can create real-time visual and acoustic environments that span physically separate areas. They can also control the recording, accessing, and replaying of the images and sounds from those environments."

Video Artists Galloway and Rabinowitz's *Hole-In-Space* (1980) is to Media Space History what "Arrival of a Train at Ciotat" is to Film History. The real-time audio-visual life-size connection between New York and Los Angeles pedestrians had people, stopping, staring, casually conversing, and gathering in awe and delight (unfortunately, no one tried to walk through the screen to the other side, but I would have tried had I been there).

Figure 5a-d. Photos of crowds gathered at both the New York and Los Angeles ends of the Hole in Space.

It also had people, as Hollan and Stornetta prescribe, meeting friends and relatives through the 'Hole' when they could have just called them on the phone. Much of the rest of Media Space history[2], however, has not been so fortunate. I believe it is because the efforts lacked temporal structure and fell victim to being appropriated as personal tools for telecommunication and video conferencing; and tools that have met limited success that have not found their way into everyday usage

2 For example, RAVE, Polyscope, Portholes, Cruiser, Montage, and Clearboard. For a full history, see W.E. MacKay, 99.

at that. Intelligent and artistic media spaces, that add a sense of awareness and a sense of play respectively to our spatial and temporal perception have proved more effective temporal 'media spaces,' though not always intended as such.

+ Artistic Media Spaces and Responsive Environments

Dan Graham's video installation *Time Delay Room* (Levin, Frohne, Weibel, 2002), in which the audience in room A could see those in room B both live and on an 8 second delay[3] on the two monitors and vice versa. The two rooms were connected by a single corridor and if one looked closely into the monitor, they could see the monitors present in the opposite room. The rooms were constructed so that spatial and temporal distances correspond, so if a person travelled from room A to room B, they would see their residual image from room A on the monitor in room B.

Time, this is what is central to video, it is not seeing as its etymological roots imply. Video's intrinsic principle is feedback.

--Gary Hill

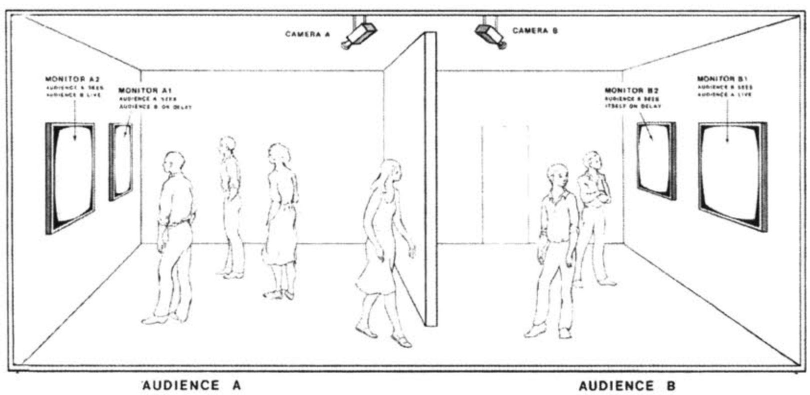


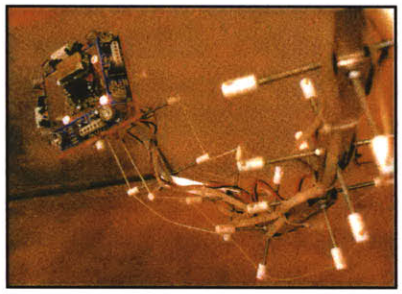
Figure 6.
Sketch of the layout for Dan Graham's *Time Delay Room* at ZKM, Germany.

The individual experienced both the past and the present simultaneously, which Graham himself says "creates both a process of continuous learning and the subjective sense of an endlessly extendible present time in flux, but without a fixed future or past states." Graham effectively created Bergson's crystal-image (Bergson, 1990) which fuses direct perception with the deferred effects of memory.

Harco Haagsma, too, explored the interplay between looking and being looked at, between perceiving and reflecting. His piece, *Biological System: Vilno* (Levin, Frohne, Weibel, 2002), however not only gave feedback to the audience as with Graham's work, but it also directly responded to their presence. Vilno was ascribed traits of a living creature with its exposed corpus hanging from the ceiling, ending with a "head" that consisted of an exposed camera and sensors. A person enters the room,

3. The time span attributed to short term memory.

and Vilno's sensors detect them, turns towards them and follows their movements, all of which is displayed on the four monitors in the four corners of the room. Through Vilno, the camera becomes part of the social environment, seeking closeness, contact, and dialogue, owing to the ability of the camera to simultaneously document and interact with the scene and action.



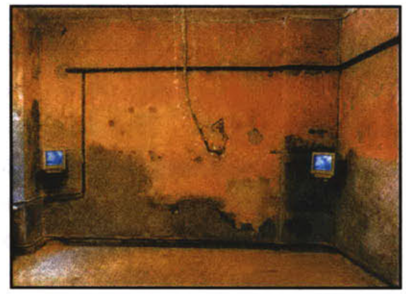


Figure 7a-b.
Biological System: Vilno
7a) The exposed camera and armature.
7b) Installation view with displays at both corners.

More recently, we have the *Palimpsest* system (Agamanolis, Piper, 2002), developed by Stefan Agamanolis and Ben Piper, which both provides feedback and reacts to the individual with which it interacts. *Palimpsest* used segmentation to detect and extract passersby in a space at Media Lab Europe from their background. Their images were then layered to form a visual that looped with a slight delay so that individuals could interact both with oneself and with other passers-by from earlier points in time that also served as a compression of the recent social happenings of the space. The dynamic video was successful in raising the awareness of the short term social history of the space, using only physical presence as the cue for image capture and response.

Figure 8a-b.
Palimpsest
8a) Installation at Media Lab Europe.
8b) *Palimpsest* after time, with delayed images of the user.

The successes of these projects have been described, but in light of the goals of this project, we see that the experiences that they create are primarily for single use. They are presented in a gallery-like setting where there generally isn't a repeated audience, and although the meaning and reflection would persist, the novelty of the interaction would not as we see in Dan Graham's *Yesterday/Today* and Haagsma's *How the System Works 4*. The prescribed use is one-dimensional because it can be, and accordingly does not need consider a fluctuating social environment. Likewise, the users are provided with little or no control over the method or recording or playback, other than their choice to be recorded or not by the constantly recording cameras. Questions naturally arise as to how such experiences can be broadened over a larger time scale in the context of more complex interactions in a perhaps less dynamic environment.

Figure 9.
Haagsma's *How the System Works*, temporarily installed in a Boys' Bathroom.

+ Intelligent Media Spaces and Responsive Environments

I began to find answers to such questions upon my arrival at the Media Lab with my introduction to my advisor Mike Bove and his Object Based Media's group concept of metadata. With dataveillance already upon us, I do not think Baudrillard is entirely mistaken in saying that old forms of representation are losing their power to connect people to each other and their environment. I don't believe that the old methods have lost their power, but rather new developments in science, technology, and society have developed a new sense of identity and society that the forms of visual communication and expression need to augment. We are not abandoning our old tools, but rather finding ourselves in need of more powerful tool. In the words of Mike Bove, "to enable this to happen, we should focus on using content descriptions in terms of objects and procedural metadata telling how to assemble the objects. Part of our research looks at ways of "understanding" content as part of its capture, and another part looks at the kinds of tools and content that such representations enable." (Bove, Jr. 2003)

This line of thinking is present elsewhere within the Media Lab. Nitin Sawhney's *Awareness Community Portals*, Karrie Karahalios' *Telemurals*, and collaborative efforts like *Smart Rooms* all demonstrate an understanding of some aspect of the user's state in the context of their environment and both display and respond to this information.

Nitin Sawhney's *Awareness Community Portals* (Sawhney, 2000) used visual sensing to provide relevant information and construct traces of people's activity over time. The 'relevant information' chosen was a grouping of an individual's photo with web-based Information they had accessed, so that an individual could learn who had been looking at what on the ever-popular Slashdot.

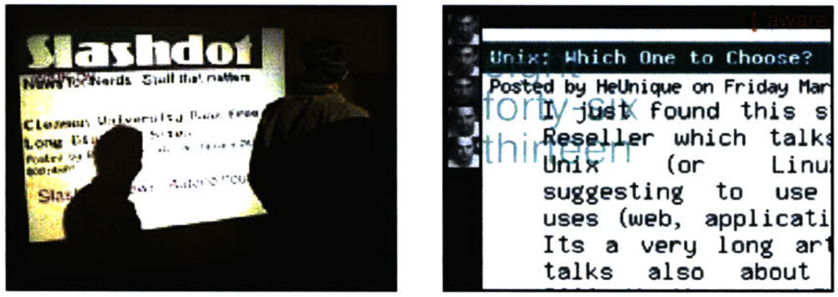


Figure 10a-b.
Awareness Community Portals
10a) Installation View
10b) Content view of *Portals* with face shots on the left and Slashdot content in the main frame.

The intent in presenting this information was to ease the difficulty of maintaining awareness and social contact with other researchers that work in asynchronous times and distributed locations. Understanding of context was also necessitated because "designing graceful shared systems that coexist with the environment requires a means to detect when an appropriate interruption is meaningful." (Sawhney, 2000) This seems to me the more difficult research problem, as social significance is not something that can be reliably detected as volume level or physical proximity can. Perhaps some combination can prove an efficient heuristic, but events don't have a set meaning value - sometimes it may be intrinsic and other times it is something that develops over time. [It is believed by this author that such a hypothetical 'social significance' detector would necessitate facilities no less than human.]

How then, can we use the variables that we can sense to constitute a more complex understanding? We see both Google and Blogdex accomplish this without any hi-tech sensing mechanism. The cataloging of what people search for and choose to share with others proves telling of a current social climate, but also reminds us, as stated above, that significance develops over time. This is somewhat self-propelled with Google and Blogdex, as the higher hit rate makes an individual more likely to click there and add to that rate. Alternatively, a user could use this knowledge to make a decision to not click there. Their transparent "sensing" methodology allows a person to see their individual effect on the system and use it accordingly.

Sometimes a simple sensing method can prove significant in our environment as well, as witnessed with *Telemurals* (Karahalios, 2002). Developed by Karrie Karahalios at the MIT Media Lab, this audio-visual communication system reinforces a sense of involvement by providing it with some intelligence to modify its visual space according to certain movements and speech inflections. The images are both rendered non-photorealistically according to pitch and volume and words spoken in both spaces are converted to text and rendered on the screen in blocks that fade away over time.

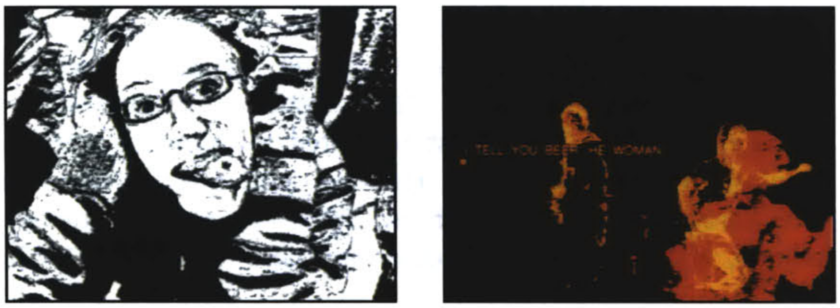


Figure 11a-b. *Telemurals* by Karrie Karahalios. 11a) Example of the type of non-photorealistic representation used. 11b) Display with speech converted to text.

The visual feedback also makes the user aware of the mechanism behind the intelligence. This awareness of their effect, no doubt affects the continued use of the system, and the feedback can be used both unintentionally (a natural raise of pitch) or intentionally as a tool for purposeful expression (for example to emphasize something the user knows the system would not itself emphasize), giving the user a more enhanced control over their representation.

The Vision and Modelling Group, active at the MIT Media Lab from 1987 – 1995, worked on developing methods for acquiring contextual intelligence for the development of what they called *Smart Rooms*. They employed cameras, microphones, and a variety of other sensors, and used these inputs to try to interpret what people are doing in order to help them in whatever they may be doing – be it looking for their keys or having a meeting. Computer Vision techniques have continued to develop in the hopes of creating other such applications[4], and the wealth of developments bring rise to the question as to why devices for media capture have not advanced in parallel. As Marc Davis of Berkeley's Garage Cinema Group states, "While the devices for media capture have advanced from mechanical to computational since the invention of photography and motion pictures in the 19th century, their underlying user interaction paradigms have

4. See http://www.nist.gov/smartspace/resources/

remained largely unchanged. Current interaction techniques for media capture do not leverage computation to solve key problems: the skill required to capture high quality media assets; the effort required to select usable assets from captured assets; and the lack of metadata describing the content and structure of media assets that could enable them to be retrieved and (re)used."

+ Video Databasing

With the rapid generation of digital information, particularly video, and the lack of devices that self-organize, much effort has been put forth to creating databases as a tool for efficiently categorizing these media, as its value often depends on how easy it can be found, retrieved, accessed and filtered and managed. "The question of identifying and managing content is not just restricted to database retrieval applications such as digital libraries, but extends to areas like broadcast channel selection, semi-automatic multimedia editing, and multimedia directory services." (Martinez, 2002).

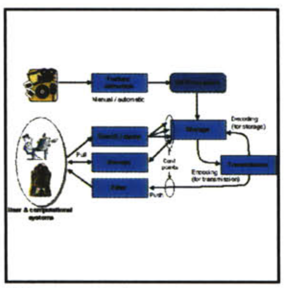


Figure 12.
MPEG-7 concept diagram of storage, retrieval, and interaction.

In October of 1996, MPEG (Moving Picture Experts Group) began development on their MPEG-7 standard as a solution to that problem. MPEG-7, formally named "Multimedia Content Description Interface", is a standard for describing the multimedia content data that supports some degree of interpretation of the information's meaning, which can be passed onto, or accessed by, a device or a computer code. "MPEG-7 offers a comprehensive set of audiovisual description tools (the metadata elements and their structure and relationships) to create descriptions which form the basis for applications in need of such information.

Figure 13.
Sports highlights on demand, made possible with ShortSports.

In January of 2003, the mgrok Corporation announced the release of ShortSports (Bergstrom, 2003). ShortSports is "a software package capable of automatically generating the highlights reels of sports games captured from broadcast television." The software uses a patent pending contextual content analysis and compression technology that was borne out of the Visual Modelling group at MIT. It allows consumers to view "entire" sports games (i.e. the highlights) in a chosen fraction of the time it takes to watch the original TV broadcast, owing to its ability to detect the interesting parts and the boring parts and use this to smoothly change the playback rate so that boring parts are skipped while important parts are played at normal speed.

Both the MPEG standard and emerging applications such as ShortSports are positioning themselves to be the Google of the media world or subcategory there of, and a needed one at that. Its technologies support a broad variety of applications requiring advanced search techniques. Neither of these techniques, however, do anything to ease the search by reducing the search pool in appropriating similar "search" methods at the time of capture. In general, say, if one is looking for information in a library, the more books that it has, it is agreed, the better chance there is of you finding what you are looking for. But if you are doing a paper on neo-gothic buildings, sometimes it is more advantageous to conduct your search at a more specialized library—the Architecture Library, for example. The distinction is between the books chosen for inclusion in the creation of the library. I believe the same holds true for specific multi-media applications. Generalized searches, equipped with tools like the MPEG-7 standard, accomplish the task at hand, but can be better put to use if they are used in coordination with other descriptive intelligence-for example, using sensing technologies to determine when to create recorded data-that limit the breadth of the library before it is searched.

The "interactive video database" project out of the Human Interface Engineering Laboratory in the Department of Multimedia Engineering at the Graduate School of Information Science and Technology at Osaka University was concerned with reflecting the intention of a user in data retrieval from a sizeable database. Their retrieval technique for video images was based on users gestures, claiming that it was a more intuitive input for spatio-temporal information. This technique can also be used as a video creation system with such gestural user interaction. They developed the flow chart for the input and output of such a video database.

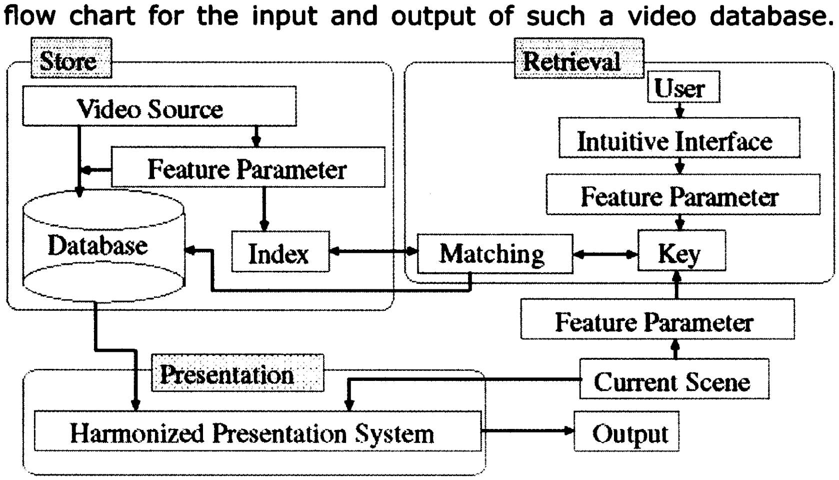


Figure 14.
Flowchart for the Interactive Video Database, showing how the "feature parameter" is used in both storage and retrieval.

In addition to the matching of the feature parameter of both the user and of the video images, they cite two other requirements of an application utilizing their model: "fusible presentation with the retrieved result and presented objects, and real-time presentation of the retrieval results." Their "interactive digital fishtank" is one example application of the proposed technique. A virtual environment in which live video and real fish are displayed together with virtual fish, it allows the user to manipulate and interact with the real fish and have this, in turn, affect the motions of the virtual fish.

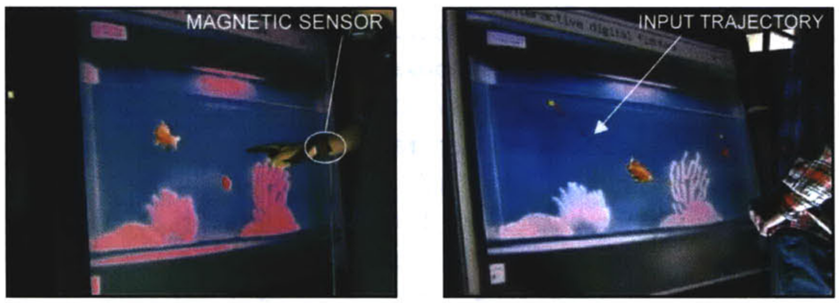


Figure 15a-b.
The virtual fishtank. 15a) The user wears a magnetic sensor so their gestures can be recognized. 15b) Those gestures are translated to an input trajectory used to retrieve clips from the video database.

This project shows us the potential power of coupling image processing with sensory "feature" parameters to guide video database storage and query. It provides a [solid] method for creating dynamic media content that changes in real time according to these sensory inputs, and gives motivations for similar endeavors that allow for a personalized and context-specific video generation.

Figure 16.
Screenshot from *Viper* showing clips (bottom) and annotations and editin guidelines (top right).

Viper, developed at the MIT Media Lab, is an example of such a context-specific video generation tool. It allows video producers to create responsive programs whose editing can change during viewing in response to preferences or profile information, presentation equipment, or real-time sensor feedback. Unlike traditional editing systems in which producers build a single linear video program, those using Viper create a database of annotated video and audio clips and other media objects along with a set of editing guidelines that describe which bits of this source material should be included and how they should be layered and assembled into complete programs for different viewing situations (Agamanolis, 2001), that could be anything from personalized advertisements to culture-specific documentaries.

Parallel efforts like ShortSports (which also served the need for personalized content presentation), the Interactive Video Database, and Viper all present different models whose components are integratable for organizing the large amount of audio and visual information that provides for both an easier and more effective means for employing this content. The next section deals with the next logical step - how this information can later then be presented.

+ Architectural Interfaces

Architectural codes both determine and reflect the social order of both public and private space and the sense of both self and community. These codes have undergone much change with the introduction of media spaces and other technology mediation. This may be perhaps because video in architectural space functions semiotically as window and mirror simultaneously, but subvert the functions of both. A mirror's image optically shifts according to a human observer's movements, varying as a function of their position in the facing space. A video image on a monitor has a static perspective that is independent of the viewer's position. As Dan Graham states, "The mirrors image connects subjectivity with the perceiver's time and space axis," where as video can be transmitted and played in a remote time and space. Windows traditionally function in architecture to mediate between private (i.e. inside) and public (i.e. outside) space so that people on either side can see into the side that they are not on.

Diller + Scofidio theorize that "The technology [of glass] that initially promised disclosure could also be availed to display false appearances, the technology that once offered democratic visibility to everyone also possessed surveillance capabilities that could look back." (Levin, Frohne, Weibel, 2002) This overexposure led to a complacency and blindness (and perhaps gave rise to paranoid anxieties that led to the development of defensive glass technologies—"privacy glass" that is one way transparent, the other way reflective). This blindness, they hypothesize was a new kind of blindness that was cured with a new hypersightedness—"a hypersightedness guided by revised hierarchies and attracted to new stimuli evolving from new strategies of display." One such strategy is their plan for the façade of San Francisco's Moscone Center, *Facsimile*. *Facsimile* is a 16 ft. by 27ft monitor that is suspended by a vertical armature that rides on a track along the surface of the Moscone's exterior, while a camera follows on the inside at the same speed. On the life-size publicly accessible monitor are live visual accounts of events within the private space of the Moscone Center.

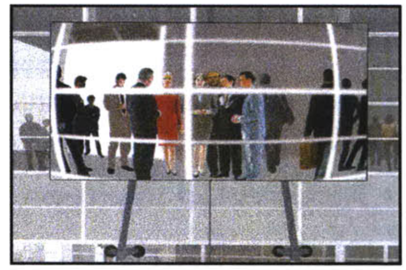


Figure 17a-b. *Facsimile.* 17a) View from the courtyard of the Moscone Center, drawing by Diller + Scofidio. 17b) Simulation of the content on the window that traverses the exterio, by Diller + Scofidio.

Media artist Frank Fietzek uses a similar user-controllable interface in his *Tafel* (Chalkboard) (Schwartz, 1997). The viewer can move the monitor both horizontally and vertically to track and scan the surface of a smudged chalkboard. In doing so, the monitor reveals texts and images once present but no longer visible on the actual blackboard. The interface effectively accumulates and conveys the palimpsestic chalkboard's information by using the images and display device as a new form of mediated recollection by merging the physical controls manipulated in the present with the virtual images of the past.

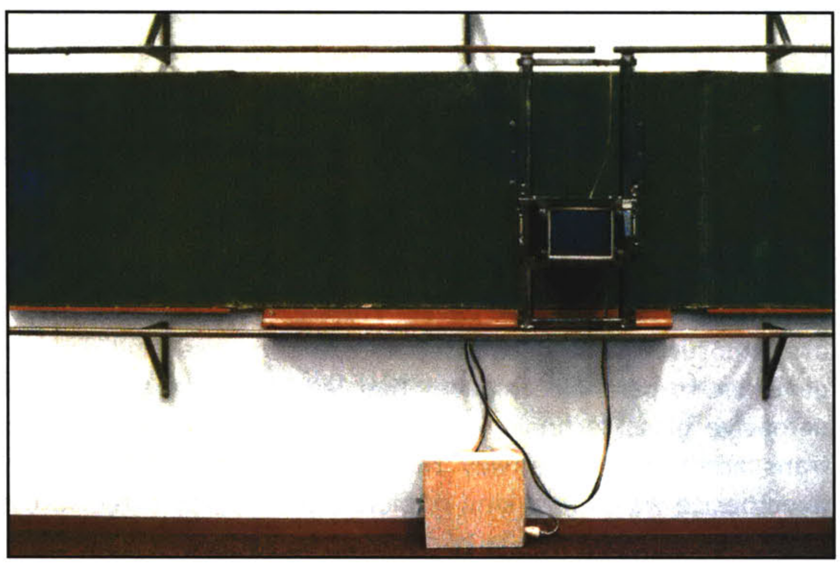


Figure 18. Frank Fietzek's *Tafel.* The monitor displays chalk writings that are no longer present.

This type of user-controllable interface is taken one step further in Jeffrey Shaw's *The Golden Calf* (Shaw, 2003). On "display" as ZKM in Germany, an object in real space—an empty plinth—

becomes the location and ground for a synthetic sculptural object in electronic space—the Golden Calf. The sculpture is only viewable through a hand-held LCD display with a spatial tracking system. As it is manipulated in front of the plinth, the virtual perspectives correlate to the real depth, distance, and reflective views the viewer has of the plinth. The calf has shiny skin, and the viewer can see reflections in it of the actual gallery space around the installation. This was achieved through photographs of the area shot earlier and digitized to create an environment that is "reflection mapped" in real time onto the calf. In this way, the real experience of the piece can only be seen by using a physical interface that lets us view the virtual, i.e. the real is dependent on the virtual.

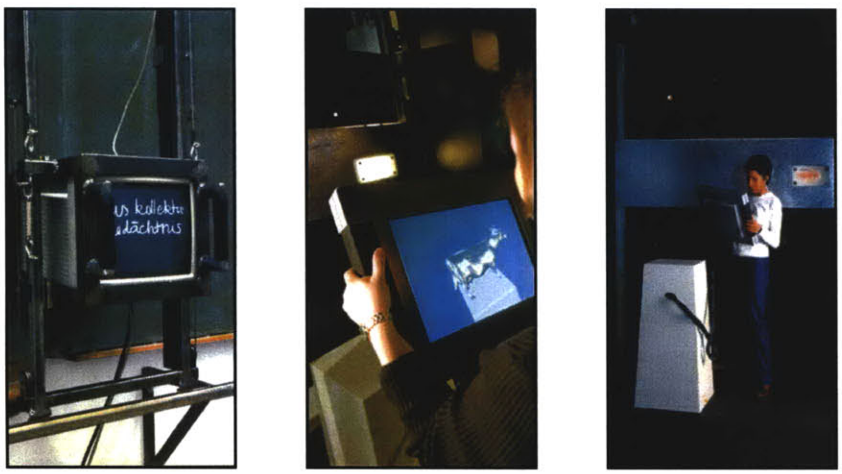


Figure 19a-c.
The metaphor in *Tafel* adapted to 3-dimensions in Jeffrey Shaw's *Golden Calf.* 19a) Interface for *Tafel*. 19b) Interface for *The Golden Calf*. 19c) User exploring the plinth in three dimensions.

P R E L I M I N A R Y   W O R K

PRELIMINARY WORK

In addition to work that is being developed around me, the work that I have been a part of as both a student and researcher at the Media Lab has been an integral component of this research. The projects were chosen to be included in this section because they were beginner attempts to illustrate the concepts and information set forth, they gave me practical skills that were needed as groundwork for this project, or they were preludes to Simulacrum in its final instantiation. For these reasons, they affected both my design implementation and goals by giving me knowledge that boiled down conceptual goals, defined anticipated technical shortcomings, and isolated desires and concerns of potential users.

+ AudioPrint

AudioPrint was borne out of Hiroshi Ishii's Tangible Interfaces class. Although the primary focus of the class is on extending users interactions with digital information beyond the traditional GUI by giving physical form to digital information, the subject matter encompasses the study of how people interact with objects and their physical environment--the physical environment including architectural surfaces like walls, doors, and windows; physical objects, like books, musical instruments, and furniture; and finally what he terms ambient media, which includes light, sound, and motion. While enhancing graspable objects and augmenting surfaces exploits the human sense of touch and kinesthesia to convey information, ambient media displays lay in the background and "communicate digitally-mediated senses of presence and activity at the periphery of human awareness." (Ishii, 2002)

AudioPrint is a sound installation experiment in physically manifesting information about the social make up of a space through sound. Simple and textural sound samples were generated using individual characteristics of an individual as they entered an enclosed space were mixed and layered on top of each other so that the ambient music was simultaneously a product of and instantiation of who was situated in the room. When the individual left, their "audioprint" lingered for a given amount of time and eventually faded, much like the conversation topics they introduced most likely did. The project was in ways a theoretical synthesizer for a music piece like Brian Eno's ambient "Music for Airports." It is intended to be in a physical place of cultural significance by way of the variety of

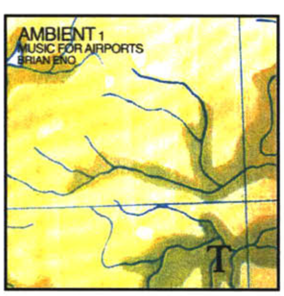


Figure 20.
Brian Eno's Ambient: Music for Airports.

individuals and social interactions that take place there, some of which are rare and some of which occur as daily rhythm.

Figure 21a-d.
Hand geometry provided the unique variables needed to generate an individual's "audioprint."

Unlike an encoded music cd that you can take with you to a train station, a hospital waiting room or your office and listen to it in the morning or at night, AudioPrint was experienced in and reflective of a single place at a single point in time. I found myself wanting to sample the music from a half hour ago when a room had just been completely vacated, because the knowledge of who had been there or what had happened in it when I wasn't there contributes equally, if not more, to my awareness of the space.

+ Cinevator

Reflective shared physical spaces was the subject of much of my work in Professor Ishii's class. A conceptual project, Cinevator, was to be an elevator video installation. Elevators move people from one physical location to another. I thought it would be interesting to build an elevator that also moved people along a collaborative, open ended story. In an elevator, people press buttons to travel to different floors. I proposed that this input could also be abstracted into data inputs for synthesizing a cinematic experience.

Figure 22a-b.
Example clips from the "Sin City" version of Cinevator. 22a) "Do not covet thy neighbor's wife." 22b) "Honor the Sabbath Day."

I proposed a fictional 10-story building, "Sin City," where each floor was assigned to one of the Judea-Christian Ten Commandments. As the elevator moves to each floor, a new section of the story can be synthesized (perhaps using some combination of recorded clips from the building, or pre-recorded clips and animations) which subtly or not so subtly relates to that topic. The resultant movie, played at the end of every day, is an abstract visualization of the everyday movements of people through the building elevators, a real-life group "choose your own adventure" story.

+ Reflexion

Reflexion, developed originally as Reflection of Presence by Stefan Agamanolis, V. Michael Bove Jr., and Alex Westner (Agamanolis, Westner, and Bove Jr., 1997) and continued with a new name by Stefan Agamanolis in collaboration with Cian Cullinan, used the layering of live video images from distinct remote locations as the "magic mirror" metaphor for this interpersonal video communication tool. The audio is presented live and the visual scene is composed by a system that responds to both visual and auditory cues. It exploits these cues to emphasize the center of attention in an effort to improve interaction and communication by enhancing it with the power to respond and adapt to its users (Agamanolis, Cullinan, 2002). In this sense it creates a type of interaction that is in some way richer than a face to face engagement. Shown below is a three person conversation against a chosen mountain background, where the varying opacity of the segmented individual is a reflective of their level of participation in the conversation.

Figure 23a-b. Two screenshots from *Reflexion.*

I joined Stefan and Cian because it was an opportunity to familiarize myself with the Isis environment, the language in which Reflexion was written, while helping them with testing involved before bringing it from Media Lab Europe for a Digital Life Conference at the Media Lab. They were also looking to extend it and for people with ideas and an interest in helping them do so and Stefan was preparing for a class on computationally-mediated collaborative activities, discussed below. In working with them, I got my first view at Isis and the power and limitations of its Macaroni libraries used for display composition and the segmentation functions that would later be extended by Darren Butler. Isis is a programming environment developed by Stefan Agamanolis at the MIT - Media Lab whose software libraries "strive to follow a 'multilevel' design strategy, consisting of interoperable layers that each offer a different level of abstraction. Designed to support the development of demanding multi-media applications, the language, its condensed syntax is geared to better reflect the "multi-layered, multi-sensory, and multi-person nature of responsive media" (Agamanolis, 2001). The tool one chooses for a given job is part of the job itself, and although I had not at this point officially chosen the topic of this thesis, it was in working on Reflexion and extensions of it that I determined the Isis language as suitable for interactive multi-media installations for someone with my background that did not include a depth of understanding of machine structures and computer organization necessary to combine external sensors with graphics libraries and device control.

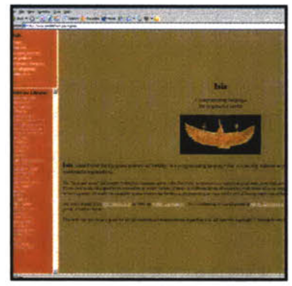


Figure 24.
Isis website: http://medialabeurope.org/isis

+ Double Exposure

My interactions with Stefan continued in his seminar co-taught with my Advisor Mike Bove on the subject of Community Maintainable Online Spaces. Among other things, their seminar explored the concepts of spatial and temporal remoteness in relation to technology-mediated collaborative activities. The course was an experiment in the subject matter itself, being held simultaneously at both the MIT and European Media Labs in Cambridge and Dublin respectively. Transatlantic collaborations were encouraged, and it was in such a collaboration with media artist Jonah Brucker Cohen that Double Exposure was developed.

The concept was to have two photo booths in two separate locations, in this case the MIT Media Lab and Media Lab Europe, that people could enter and have their picture taken and printed with someone that had simultaneously done the same at the other end of the connection, or alternatively their photo would be composed with people that have previously been in that booth. The software was implemented in Director using two CCD cameras and video capture cards at both Jonah's and my workstation, but the physical installation was never completed. This was the first of my endeavors that attempted to collapse time and space using a single visual frame.

Figure 25.
The classic photobooth experience, which was re-examined in DoubleExposure.

Figure 26a-c.
26a) Concept sketch for the version of DoubleExposure implemented. 26b) Concept sketch for the different scenarios for use, including 2 booths that are spatially remote, 2 booths that are both spatially and temporally remote, and a single booth that keeps an ongoing memory of everyone that has been through it. 26c) An example DoubleExposure filmstrip.

The use of the well-established photobooth form lends to the novelty of the experience, introducing a new element to the standard script of the photobooth experience that redefines it with a fresh technical vocabulary. But the vocabulary is specific to this application, as the concept does not extend to a less contained space, for example a public courtyard or classroom, that is used in a variety of ways that has more complicated functions. Such a system would need to be general enough to encompass all such uses and functions, but reactive enough so that an individual could sense their effect on it.

+ MotoFoto

My first experience in exploring methods to both sense and react to a user in this manner came with my experimentation with the Cricket, and later the Tower, developed within the Media Lab in the Lifelong Kindergarten and Grassroots Invention Groups respectively. Both the Cricket and the Tower are modular development systems for designing and prototyping computational devices. Physically, the Tower consists of a primary foundation layer with a central processor equipped with a PIC16F877 microcontroller from Microchip Inc.™ Additional modules can be added to the stack to perform sensing, actuation, and data storage (Lyon, 2003). One such project, entitled MotoFoto, used a single proximity sensor and four Lego motors to actuate a flip book at a speed that varied with the "user's" proximity to the structure. It was an attempt to have a single story morph its presentation based on an individual that was present that may or may not be actively using the system. As the use came closer, the speed of the motors would slow as would the story, as if the physical closeness correlated to the closeness of his inspection. When there was no one within close proximity, the motors would whirr away; trying to give the complete story to people that glanced at it as they passed by.

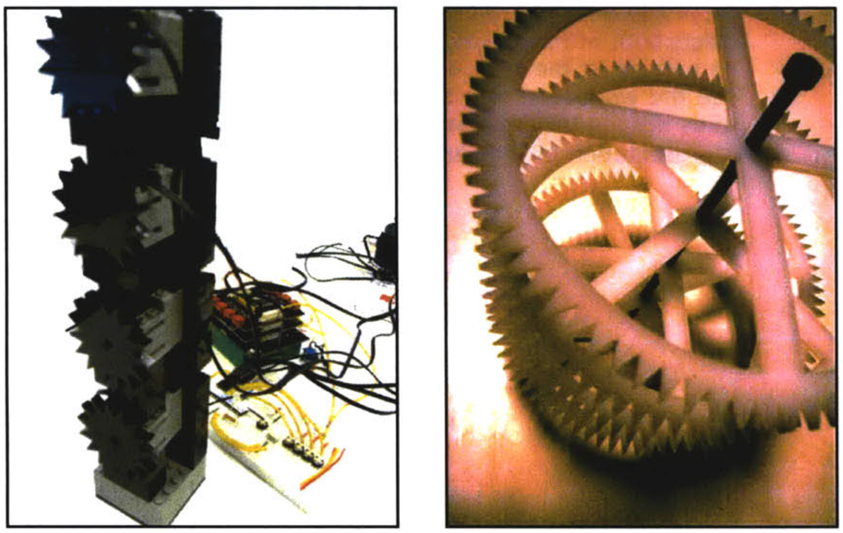


Figure 27a-b.
27a) Electronics, including 4 buttons, a proximity sensor, and 4 motors with attached gears that drive the system. 27b) The physical structure, driven by the motors, that holds the photographs.

In one way or another, I see this thesis project as a continuation of each and all of these projects. While AudioPrint successfully manifested information about the social make up of a space, this information was not accessible to someone that was asynchronously present - i.e. there was no memory of it. Cinevator used abstracted data input to dynamically edit stored video clips, but in a "use once and destroy" manner that needed to be re-edited on a regular basis to remain interesting. The applicative usefulness of audio and visual cues was proved with Reflexion, which also introduced the concept of traces alongside Double Exposure which, to me, was a key concept and visualization method for bridging the asynchronous gap. These ideas coupled with the knowledge of basic sensor knowledge that allows me to define the responsiveness of the responsive environment.

Figure 28.
Magritte's *The Human Condition.*

Simulacrum is a socially responsive environment that manifested events and happenings from that space. This information is accessed asynchronously and presented in a matter that dynamically reflects the present happenings in the space. It puts the past into the present. Physically speaking, it is an controllable window into the past. The fitting interface to me was a user-controllable, free control monitor not unlike that seen in *The Golden Calf*. The underlying metaphor not that different from Magritte's *The Human Condition*, of which he says:

> "In front of a window, viewed from inside a room, I placed a painting which depicts precisely that section of the landscape obstructed by the painting: that is, the tree of the painting concealed the tree located behind it outside the room. In the mind of the viewer, the tree exists both within the room, i.e ., in the painting, as well as outside in the real landscape . And that is how we see the world: we see it as something located outside of us despite the fact that what we see is only a mental representation of that which we experience within us."

I M P L E M E N T A T I O N

IMPLEMENTATION

+ System Overview

This project aimed to be a new form of technology mediated communication that allows people, whether office mates or strangers, to see and interact with each other independent of the time continuum by situating this representational social memory into the environment. Media spaces generally serve to collapse spatial gaps by providing a synchronous audio and video connection between two geographically remote locations, as mentioned in the Background section of this document. This work aimed to collapse *temporal* gaps by providing an asynchronous but continuous audio-visual presentation of traces from the past (ranging from days to seconds ago) from and in a set physical location. Being a physical installation that is continuously running, it also affords connectivity between spatially remote participants.

Figure 29.
Photograph of the Garden area of the MIT Media Lab with an early version of the Simulacrum system.

The MIT Media Lab consists of both shared offices and shared work areas that are common to multiple research groups. The lab thrives on the interaction among the members of research groups to maintain a dynamic research and social community. The activities and personalities that inhabit the Garden are as diverse as the research that goes on there. Working asynchronously and remotely are part of the daily rituals of the inhabitants, and two people can go days without seeing each other or knowing what the other has been up to. Although this is not a direct threat to the functioning of the lab, labmates' distance--whether spatial or temporal—can be bridged with technological embellishment that serves as a collective memory for the members of the community. Aside from the inhabitants, the Garden consistently draws a large volume of outside visitors unfamiliar with its population, climate, and rituals, who can make use of such a device, considering that they are usually seeking, among other things, information on the general happenings of the lab. For these reasons, along with the familiarity and convenience of being my personal workspace, the Garden was chosen as a suitable environment in which to explore this project.

Situated in the Garden, the Simulacrum system is a projected video display intended to run continuously, composed of five independent software programs running simultaneously on a single PC, two wide angle CCD cameras, a projector and a screen, a microphone and speaker, three buttons, a proximity sensor, a potentiometer, and associated Tower and PIC16F877 microcontroller circuitry connected to the PC using the RS232 serial protocol.

The projection displays a user or group of users inside pre-recorded video footage of events, or "highlights" in that space in the past composited with a current view of the present. With a user's input, the display can change to facilitate the manual search of the database. Though the screen was initially intended to be placed in front of the capture area of the cameras so that the video image covered the actual area, not unlike Magritte's *The Human Condition*, space constraints led to the current set up where the screen functions as a mirror, rather than a window, onto the space. When used as "mirror", the users can see themselves as part of and elastic timeline of images from the past. The signal level of the embedded proximity sensor, microphone and movement within the frame of the captured images is used to control the sequence and speed of that timeline automatically. Additions to the database are also made automatically using the same meta (sensor) data that is used to dynamically edit the presentation of the video images. This can be manipulated directly with intuitive controls that allow users to change these editing variables as well as browse the database and manually record a happening that is then added to the database.

An addition to the database is not simply the video--a time-stamp and the associated ambient sensor data is also stored with it. The database itself is dynamic too—it uses the same rules and

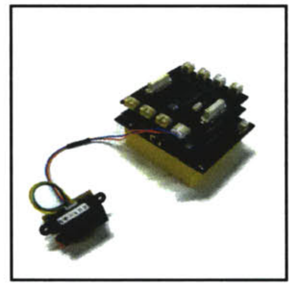


Figure 30a-c.
Some of the components involved in the system. 30a) Tower and proximity sensor. 30b) One of two CCD cameras. 30c) Potentiometer and one of three buttons.

metadata for editing the database for presentation to regularly compresses and overwrite itself. This is part of its elasticity—the current display is a variable compression of the database, but the database continually compresses itself, so that events that took place further away in time become more and more compressed while recent events are more fully displayed, and more significant events are retained for longer. The augmented space containing the system becomes an evolving artifact of the community and a member of the community itself, owing to its ability to store this information created by the individual and the community and to use that information to interact with them.

The active capture, reactive display, control interface, video keying and database management is accomplished with five modular Isis programs that communicate using UDP signals within the computer as shown below:

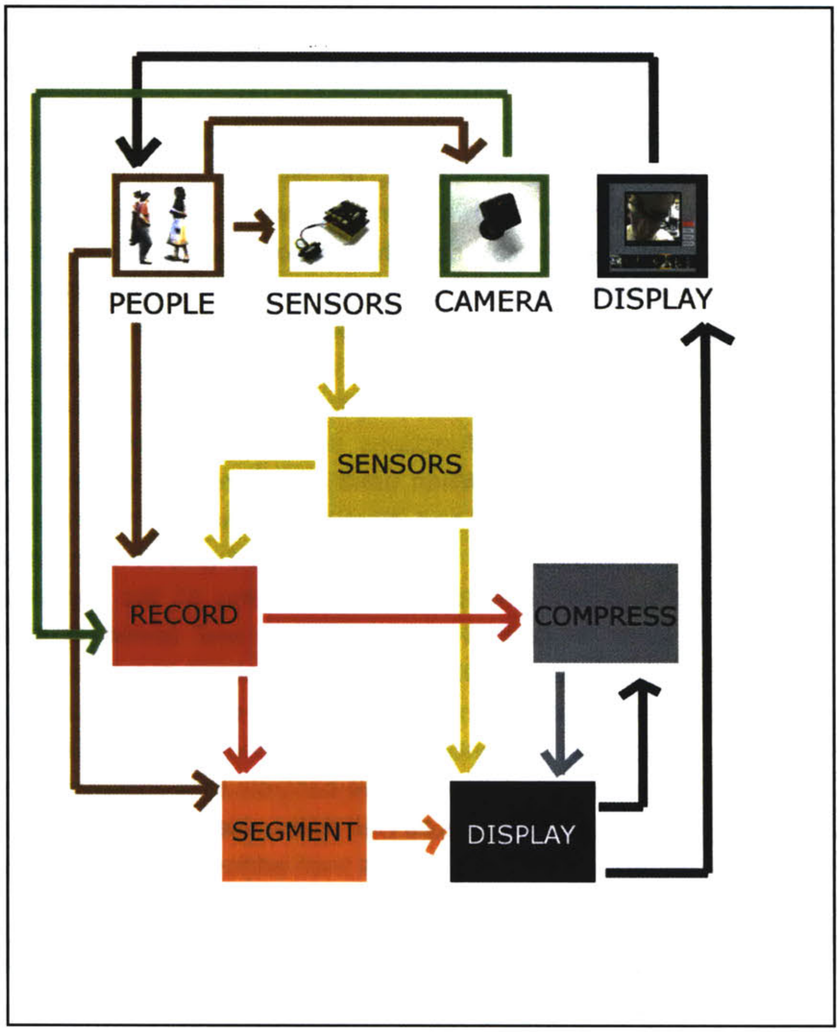


Figure 31. System Diagram of communication between modular Isis programs (colored blocks) and physical devices (top).

* Active Capture: The sensors and cameras are continuously sending their signals to the computer for processing—the sensors through the Tower connected to the computer via serial port, read and parsed using the Isis *sensors* program and the camera through the video capture card that is controlled using Isis *segmentation* and *record* programs. Communication between them allows us to record only when there is something considered worth recording.

* Reactive display: The *sensor* program also communicates with the *display* program, which uses this information to change its display dependent on the current activity in the space. It also lends to the user's sense of interactivity, as they both see themself in the foreground (accomplished with the segmentation functions) and also eventually see their activity added to the database by seeing it played in the background video loop.

* Control Interface: The *sensor* program also parses the information from the physical controls from the Tower—the three buttons and single knob—and sends this information via UDP to *display* program which controls the graphical interface for the database search, can communicate to the *record* program and initiate capture, and can change the settings for the real-time editing.

* Database Management: The *display* and *compress* programs share the duties of the database management. The *display* program looks for signals from the active recording to detect when new material needs to be added to the database and put in the display loop according to its associated data. It also sends a signal to the *compress* program when it is oversaturated with such material, where it is condensed.

One great advantage of implementing communication between the program's modules using UDP signals is that each single program or any combination or programs can run on a separate computer, which leaves the system open to reconfiguration. At the beginning of this section, I noted that this project connects people that are both temporally and spatially disconnected. In its current instantiation, these users still need to be somewhat co-located (i.e. they both need to be in the Media Lab Garden at some point in time), but it could just as easily be implemented with the capture and display in two geographically separate locations, so that viewers in one (or both) location(s) can see a collapsed history of activity at the other "capture" location. Alternatively, there could be multiple complete stations in a single space, sensing from a given space could affect the recording and display in another location, or the present view of one space could be composited with the past view of another, or any number of other configurations.

+ Sensors

Guy DeBord says on the construction of situations, "A person's life is a succession of fortuitous situations, and even if none of them is exactly the same as another the immense majority of them are so undifferentiated and so dull that they give a perfect impression of sameness." (DeBord, 1967) It is the rare social moments that stand apart from this sameness that I intended to capture. Sensor technologies can "perceive" human presence and activity which largely contribute to social conditions. Optical proximity sensors are used in this project to not only detect that someone is present, but also the frequency with which people have been present in the Garden area. Observed high proximity, high frequency traffic, and large changes in proximity and traffic are considered reason to record the scene. The actual thresholds for such measurement were created by the guidelines of personal observation. That someone is present is usually of substantial interest to the onlooker (the majority of people glance at someone as they walk by), but the presence of more than one person, or continued presence of an individual is more likely to hold social value. It is using this common sense social knowledge that "scores" are given to encapsulated moments of recorded images. When this score is above a defined threshold, a video clip is recorded and added to the database. In doing so, the volume average is measured, whose value also contributes to this score. The program sends this information—the "score"--to the display program so that it can be used as a variable in the editing of the clip.

+ Segmentation

At the core of the segmentation routines is a background averaging and difference algorithms written by Darren Butler (Butler, Sridharan, Bove, Jr., 2003). It is being used to both create the icons in the lower banner menu, and also to reflect the images of people presently in the space by situating these images in the display.

Figure 32.
Screenshot illustrating the segmentation algorithm operating with different morphological operator and combinations thereof.

The icons are created by segmenting an image from the middle of the clip it represents with the current background seen by the camera. This process "detects" differences in the average visual scene than the scene in the recorded segment, which were most likely the reason that the segment was recorded. The result is an image that (highlights, accentuates) the activity of the clip.

Figure 33a-b.
33a) Image as taken from capture device. 33b) Image after background segmentation used as an icon in the menu banner.

This process is sensitive to the arrangement of the space (moving a table or chair), person's motion, and the changing of the position of the camera. Highlighting these activities brings to the foreground the particular moment that makes the clip it represents different from the rest, and eliminates the extraneous background information. In doing so, the system more clearly presents the information it contains and facilitates a users search of that information.

The mirror effect is made possible with the help of the Open CV morphological operators acting on the results of the background segmentation. An object in motion—most likely a user—can literally be cut-out of their background and pasted into a background from the past. This gives the user the feeling of being in that time as they can see themselves as the action carries out behind them in the projected image.

Figure 34a-b.
34a) Segmented image
34b) Image composited with background in the video scene.

+ Record

Figure 35.
Screenshot showing record indicator (red rectangle, middle right).

The record program is fairly straightforward. It can receive a signal from both the sensor and display programs, and when it does it carries out the record command according to variables set for the program. While recording, a visual cue—a red rectangle—is displayed so that inhabitants are aware that they are being recorded. The record variables--duration, fps, brightness, contrast, quality, audio frequency and bitrate—can be changed at startup. The current duration for a single clip is 10 seconds at 30 fps. At this time, the sensor variables are not used to variably change the video clip at the time of recording, though the framework is there for that to happen if desired.

Just as the display can go into and ambient mode, the recording process can go into a similar idle mode. When the activity in the Garden tapers off, as measured by the time since the last recorded event, the recording is reduced to one frame every 5 minutes (this value was chosen by the author and can easily be changed). The idea behind this is so that periods of little activity are not completely cut out of the history, they remain, albeit shortened, to preserve the continuity of time in the presentation.

+ Interface

The physical interface for the Simulacrum project consists of three buttons and one knob. Though a person's presence and action (i.e. their body) is the seamless interface to the system, the buttons and knob are the method by which a user can explicitly communicate with and alter the state of the system. Two of the buttons are used to toggle the cursor left or right through the menu of icons always placed at the bottom of the screen. The third button is used to explicitly record an event. This button was included because, although the system has intelligence, sometimes a human is more capable of detecting when something is of social interest. It also affords the user some room to use it for reasons other than its expressly intended purpose. The knob can be turned left or right to slow down or speed up, respectively, the speed of the video playback loop displayed.

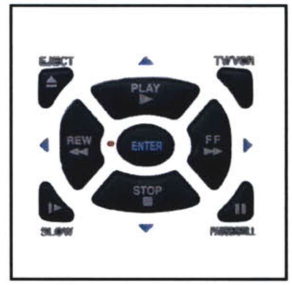


Figure 36.
Typical VCR control.

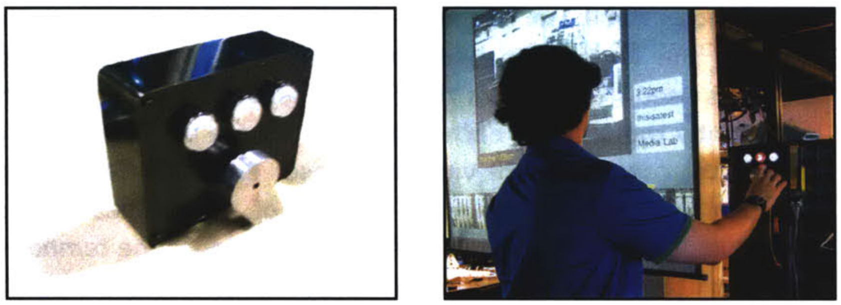


Figure 37a-b.
Control Interface. 37a) The physical elements. 37b) In use, lighted, labeled and wired up.

The interface is not unlike that of a VCR, and even uses the symbols associated with typical audio-visual control devices. Though this was not the original intent, for a target audience familiar with such devices it is intuitive and can be used as intended without any instruction.

+ Display

The display program is the nucleus of all of the other programs. It is also the program that most directly communicates with the user. There are two modes of display: active and ambient.

Figure 38a-b.
Active and ambient modes of display. 38a) Active mode showing yellow user-controlled cursor and preview window. 38b) Ambient mode showing full screen video and automated yellow cursor indicating the time of the current clip playing. Also shown is the segmentation of the present view.

The display is "active" when a user is using the physical interface to view individual clips. The filmstrip menu across the bottom of the frame displays icons of the individual clips described above. The user can move the yellow cursor across this strip and in doing so, the movie selected immediately begins to play in the preview window. In this mode, the movie is accompanied by the original accompanying audio. A caption containing the time of the movie event and the present time are both presented in the forefront of the display as shown in Figure X. When the user becomes inactive, the movie that he/she was last viewing loops until a set amount of time has passed (presently set to 30 seconds), the display shifts into its ambient mode.

It is in "ambient" mode that all of the aforementioned re-editing of the video database happens. If the speed is not explicitly set with the knob by a user, the median of possible editing speeds is used. This speed, the current sensor values, and most importantly the score of each individual clip are used to variably present the contents of the video database. Presently, the editing is accomplished by selecting a fraction of the frames of the entire clip; that is, the higher the score, the higher the *temporal* resolution that activity has in presentation. It should be noted that it is at this stage in the editing process the individual images could be edited for a non-photorealistic presentation. This was implemented, but in doing so, the system no longer responded in real time, so that feature was removed. Possibilities for resolving this are discussed in the Project Evaluation section of this document. In this mode the cursor scrolls across the menu bar as content from each of the video clips is being displayed.

The system is designed so that the display is most often in this ambient mode. In this sense it becomes an ambient display—the sound is muted and the video is displayed at full-screen resolution. In this mode the user sees themselves live in the projection. The name ambient, however, should not imply that the system is not active. A user can see both themselves live, and if they create a sensor reaction, a signal will still be sent to record and that clip appropriately edited into the loop playing. A user can also use the physical interface—the record button--to initiate a record or they can use the knob to change the playback speed of the edited clips. Use of two (left and right) toggle buttons used to search the database will switch the user back into "active" mode.

+ Compress

Though it would be more effective if the Simulacrum system could keep a complete memory from launchtime until present, both memory constraints and, more importantly, privacy issues led to the design choice to have the database continually overwrite itself. Automated recording processes are a technology that is met with discomfort and distrust in society at large. As discussed in the introduction, appropriation of such technologies into surveillant devices threaten the individual's sense of privacy. It is not the goal of the project to use the recorded information collected to document activities with suspicion or distrust. They are collected with the mentality that they can provide valuable information that, when used collectively, dissolves the threat posed by a system that is being monitored by a single person in power inaccessible to its "users."

When the database reaches a defined limit—50 clips—this program re-edits those clips and drastically reduces the size of the database. The number of compressed clips is a fraction (whose value is dependent on the scores) of those in the original database. The limit was chosen to be at the point when the icons became unintelligible and would cease to effectively display their information.

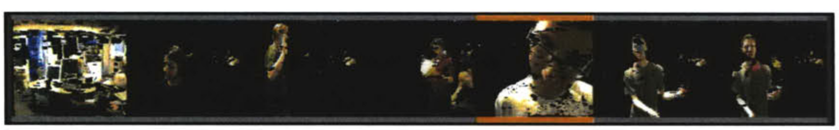


Figure 39a-b.
Xa) Icon menu bar at just under capacity. Xb) A less congested menu bar with just a handfull of movies in the database.

The compression is carried out by trimming each clip by extracting frames at a given interval and "gluing" such compressed clips into a series of large clips that match the duration and frames per second of those directly recorded. Although they match in duration and frames per second in terms of play time, they span a longer period of time and thus appear faster because of this higher time lapse. The rules for compression are similar to those for the dynamic display edits (see below). The higher the score of the clip, the more frames are preserved from this clip. The resulting amalgamation of multiple clips is given a score that is the average of all clips whose frames are included in it.

This uses a "survival of the fittest" metaphor with the exception that eventually everything that enters the database, eventually is overwritten. The resulting visualization can be described using the metaphor of the elasticity of time created by memory.

+ Settings

This system is designed to be flexible, owing to the desire that it be community-maintained (ala the Situationists, DeBord, 2002). Haven been writing in the Isis scripting language makes this task fairly simple. The *settings* file contains many of the parameters that direct the general functioning of the system and can be changed there quite easily. The user can adjust:

* The colors of the "active mode" display.
* The size of the preview window in the "active" mode.
* The size of the icons of the movie menu banner.
* The time before the system switches from "active" to "ambient"
* The time before the system begins to idly record.
* The maximum number of movies that are in the database before it compresses and overwrites itself.

E V A L U A T I O N

EVALUATION

The Simulacrum environment described above is the result of five months of research and development. While the project aimed to be a dynamic space both in its ability to document the past and it its ability to reconstruct these events in the present, as with any project, it is also conceived as a compromise between time, resources, and skills needed for its implementation.

+ Application

The system was run in the Media Lab's Garden intermittently over a two week period, leaving room to make small changes, modify the interface, and add new functionality, and has run continuously over one week during the month of July. The physical setup and projected environment proved enough to attract the attention of both regulars and passersby, owing to the fact that the moment a display device is set up, people tend to want to know what is being displayed. No users were solicited for the evaluation of this project.

Figure 40a-b. Photos of Gardeners (40a) and non-Gardeners (40b) using the system.

The target users were anyone and everyone that happened to pass through the Garden; this consisted of its inhabitants, non-Garden-members of the lab who visited whether to talk with one of its inhabitants or to attend a lab-wide meeting or tea, friends and family of its inhabitants, and outside visitors and sponsors of the lab. More attention, however, was given to those that observed and used the system over the entire three week period (i.e the Gardeners and Medialabbers). Their extended use gave me a greater ability to see how this system could affect their daily life, particularly their perception of their workplace, knowledge of the events that happen around them, and social connectedness.

++ Gardeners

The reaction to the system from the Gardners can be classified into three categories: as a filtered memory, as an interactive mirror between the past and present, and as a surveillant tool.

Those that saw it as a filtered memory tended to view it most often in its "ambient" mode, and rarely used the interface to access particular clips in their entirety. Alan, a second year graduate student whose name has been changed, commented that it was like a Tivo for Garden life, claiming that he got to see only the parts that were of interest to him. Another student, Ned, said that it fit with his newfound philosophy that the most entropic moments in life are the ones that are worth remembering. When this class of users did use the interface, it was most often to record themselves, claiming that they wanted to make sure that they got their due representation in the Garden "highlights."

Using it as an interactive mirror also bred a lot of self-recording, both by action (thus triggering the proximity sensor) or direct use of the record button on the interface. Max and David, both graduate students and the two most prominent of such users, were not concerned with making sure they were documented as Gardeners, but rather sought to interact with themselves. Accordingly, they would use the system in its "display" mode thus repeating their clip to simulate fights, discussions, and dances between their past and present selves.

Figure 41a-b.
Two examples of individuals interacting with themselves through the system.

Both expressed frustration that the clip with their dual selves wasn't kept as a record, but the rest of the community was entertained by seeing them punch the air and dance by themselves later. I also found enjoyment in this aspect of the project, most notably one day when I was standing on a table that is in the camera's view and looked to the screen to observe myself levitating in the picture. The table had apparently been moved 2 feet over and I was standing where there had previously been no surface to stand on. Thus in the resulting picture, it looked like I was standing in the air.

Figure 42.
Photos taken from Simulacrum footage of late night TV watching.

Quite a few of the members of the Garden felt that this system added to the number of ways that they could be further monitored, noting that "at least this way [they] know what [their] advisor could see." During the evaluation period, there was a series of laptop thefts on the 4th floor of the Media Lab, and one student commented that if my system had been running up there (the Garden is on the 3rd floor), we could know who the perpetrator was. Also, despite the large numbers of students that keep late hours, night time (12am - 8am) shows the least amount of activity in the Garden. During these hours, the system usually enters the idle recording mode mentioned above, where it captures a frame every five minutes, as opposed to the daytime where an average of 6 full events are typically captured. The Garden inhabitants, myself included, noticed that an unknown individual watches television nightly at 4am which was captured by the system. Their identity, despite our efforts to identify them, remains a mystery.

++ Medialabbers

Being a research lab, most Medialabbers often approach a project with questions of the underlying technologies. This makes it a good forum for receiving critical analysis, but sometimes natural responses get overlooked. Many non-Gardener members appreciated the use of segmentation both in compositing the past and the present and in highlighting the action of individual clips in the display mode, but I found it hard to determine what valuable non-technical information they left with. Mark, a PhD graduate student, upon observing one single loop of clips moaned that he knew so few people at the lab nowadays and asked me who a number of people he saw were. Perhaps he left with a greater probability of interacting with those he saw in the recorded images. At a lab-wide tea, a group of students watched the display and jokingly stated, "so this is what you guys do around here," having seen video primarily showing people dancing and making faces into the camera.

++ Friends and Family of Garden Members

The reaction from friends and family of Gardeners was uniformly consistent. They would initially play with their reflection in the display, but as they were watching themselves in the past clips, their level of interest spiked when they saw themselves juxtaposed with their friend or family member—their link to the Garden. Following this they would either await seeing the clip again, or use the interface to manually access that clip or collection of clips featuring that person that linked them to the space.

++ Outside visitors and sponsors

In observing outside visitors and sponsors of the lab interact with this technology, I found that they, too, were more interested in the application of it rather than their personal involvement with it. In this respect, most corporate R&D representatives saw it as applicable to surveillance or monitoring tools. One suggested application was to keep an ongoing log of who was physically present at *their* lab, hypothetically placing both the display and sensing mechanisms at the entrance to the building. Another application suggestion, that interestingly was initially a focus of the project, was to use it with equal emphasis on audio and video capture as keeping a detail of highlights of meetings and conversations that could take place in a hypothetical conference room, so that those that missed that meeting could use the system to get a recap of the dialogue and keep updated.

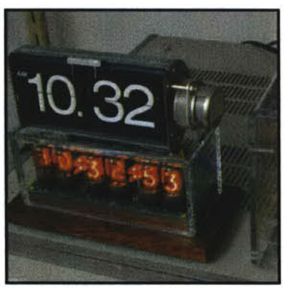
Figure 43a-c.
Some of the Garden's varied lighting conditions.

+ Project Evaluation

The current implementation of the situated memory features an optical proximity sensor and microphone as the primary means for sensing presence and activity. The Garden, like most spaces, is subject to highly varying lighting conditions. Under some varieties of these fluctuating conditions, the optical proximity performs poorly. The microphone, too has a limited range that is not large enough to span the entire area that is covered by the wide angle camera lenses. So Alan, the second year graduate student mentioned above, would be mistaken in thinking that everything of interest to him is displayed in the highlights. Many events evade the proximity sensor, and the microphone is often just out of earshot of activity in the Garden. This also says nothing of the fact that proximity and audio levels, while a start, are by no means a closed system for the social state monitoring that could be used in this system.

The program's foundation is set up so that the addition of new sensing means and methods can easily be incorporated to the current "score-keeping" method that uses proximity and volume. Use of these scores to create the weighted timeline of video material proved effective, but it was not immediately obvious to users tat this same information was also being used in real-time to further exploit that timeline. Original project goals planned to use the immediate sensor values to also manipulate the photographic presentation of the images. At present, video clips are stored using the Isis movie format which stores the individual images as JPEG files. Thus to edit a movie, one must decode the JPEG image, edit it, and then re-encode it. The time costs of decoding and encoding the image prevented the program from running in real time, and this functionality was removed. This also left the display without a definite aesthetic other than the graphical layout, as the live video was played back as it was recorded. The choice to store the data in the Isis movie format was motivated by memory capacity limitations, but given that this database is capped at 50 movies (a number that could for practical purposes be increased, but was chosen because it was the limit at which the icons were intelligible and the program could handle the editing demands), storing the images without encoding them is a viable option and would greatly reduce the time required to manipulate and display them.

With that one exception, the simple database structure proved effective for the dynamic storage and retrieval the program modules required, most especially for the combination of audio and visual information. The audio information, however, is grossly under used in the current implementation. While the 10 second audio clips paired with the movie clip proved interesting and novel, it is more suited for an audio-photography project than one in which people are intended to glean useful information from it. My lack of knowledge in manipulating audio files to provide an analogous "time-lapse" view of audio events and time limitations prevented me from exploring the possibilities that they present, as a simple speeding up of the ambient audio would prove equally insignificant.



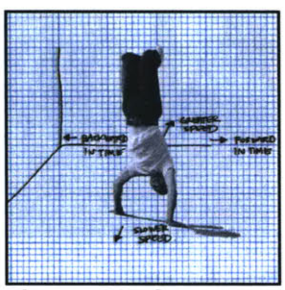


Figure 44a-b.
Some other possible interfaces. 44a) Manually controlling a clock or calendar to control place and speed of video. 44b) Movements in physical space mapped to movement in the virtual time of the video.

The simple "VCR" metaphor of the interface allows users to effectively control the system with no learning curve, but it is not the only implementable interface. The icon banner was continuously displayed with the mentality that the projection screen could be eventually be used as a touch screen, and one can imagine an interface that exploits the user's visual focus of attention, both of which would be equally intuitive, but would be even more seamless. To other extreme, the interface could also have involved interaction that supplemented the visual experience, for example one that mapped movements in space to movements in time (i.e. the video could mimic a user moving or leaning forward or backward at varying speeds) or one that mapped the manipulation of a physical calendar to the manipulation of the associated visuals from that day and time, all of which were not materialized because of time limitations.

Although the project was lacking in an aesthetic responsiveness, the compositing of present images with the past video clips proved a highly effective way to collapse temporal gaps, and give the present a role in the past. Many users responded to this effect by saying that they felt that they had "actually been there." And first time users initially do a double take over their shoulders to reassure that the people they see standing behind them on screen aren't actually physically present. Users also found they could manipulate objects in the present physical environment and give them a manifestation in the past environment. The segmentation, however, suffered from noise that could not be eliminated with the OpenCV morphological operators in place in the Isis library. The noise is most likely generated from the small cameras and could be most effectively eliminated by using a higher quality camera.

Students' comments about the laptop theft and the potential use of this system in identifying the criminal, also bring up an interesting point in the project's evaluation. If it was to be appropriated as a surveillance tool, it would be in a manner that serves the community. This is reminiscent of the lesson learned from Jamie Wagg's *Shopping Mall, 15:42:32, 12/02/93* video installation (Levin, Frohne, Weibel, 2002). He displayed images of taken from security footage depicting two boys with a toddler.

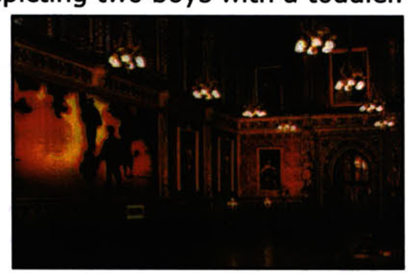


Figure 45a-b. Jamie Wagg's *Shopping Mall, 15: 42:32, 12/02/94.* 45a) Surveillance clip. 45b) Installation view.

The toddler was James Bulger; the two boys were eventually found guilty of kidnapping and killing him. The art piece was met by an infuriated press and angered Bulger family. The work was not meant as an intrusion on the private matter of the Bulger's family loss; the message was one intended for the public. One point of view is to see it a statement on the inability of surveillance cameras to prevent crime, other than for it to serve as a threat to potential wrong-doers. From another point of view, one could see it as a wake-up call showing us that the social circulation of such images is a potentially effective way to solve crimes after they have been executed. Marc Cousins states on the project that "The shot of the boys never represented the concerned look of parental care. It was the empty look, the look of a security camera. It does not look; it records." The social gaze that chastised Jamie Wagg for his piece was one of parental concern, one that could see the crime happening in this clip. Though this system does attempt to look, and not merely record, it most likely could not have prevented the theft of the laptops. It does, however, distribute viewing privileges democratically so that every member of the community has an opportunity and/or responsibility to survey and maintain their environment.

The short three week evaluation period left me unable to appropriately observe and evaluate the longer lasting effects of the introduction of this system to the Garden workspace. The generality of its application leads me to believe that it is either a) suited for a more dynamic environment where the inhabitants couldn't already guess where someone was or what they were doing, as we usually can of other Garden members based on their fairly regular routines and lifestyles, or b) needs to be appropriated for a more particular task, be it an ongoing videolog of personalities in the Garden or a tool for encapsulating the essence of a meeting that takes place there.

+ Future Directions

The construction of the basic structure of the Simulacrum system was completed for this thesis. The system shows both areas for its improvement and could stand for a variety of added functionalities that would increase the technology's usefulness. Some of these improvements and additions are listed below:

The physical interface could also provide users with the ability to maintain the database themselves, providing them with the ability to delete archives, increase and decrease the "score" of a given clip, control the compression of the database.

The interface could take a form that more actively engages the user in a manner that supplements the metaphor of the visual experience, for example allowing users to adjust the arms of a large physical clock to control both the speed and temporal index of playback.

The viewing patterns of the video clips could be incorporated into the weighting of them, as what people choose to repeatedly watch is a good indication of what people find interesting or get information from, as proved by Google.

The compositied view of past and present could be saved as part of the archive. This would introduce a third and even fourth, fifth and sixth temporal layer that could allow for users to not only see what happened in the past, but also how people have reacted and interacted with that event, making the record an even greater palimpsestic document of events.

Alternate configurations that exploit the system's UDP-afforded modularity could be experimented with, perhaps involving another display area in a spatially remote location within the Media Lab or even in a completely different geographic location.

The database structure could be used in conjunction with another architectural display form, for example a hand-held LCD screen like that in the Golden Calf, that incorporate different architectural metaphors.

Again, the images could be rendered non-photorealistically to convey information such as level of present or past activity, time elapsed since capture, or usage statistics.

Sound could be more directly presented for the linguistic information it contains, or the sensors could be use to synthesize a more ambient audio presentation, not unlike the AudioPrint project.

From a human's point of view, it is easier to declare that something is worth being forward-looking and initiating audio and visual capture at the onset of an potential event, it could function to document a moment that no one knew needed to be preserved until after its completion owing to its ability to be constantly streaming.

Though I stated above that the listed ideas, if implemented, would be an improvement or addition to the system that would increase its functionality, like the project itself, they're expected use is experimental. I think all of them are worthy extensions to the system based on my observations and interactions, as I am sure there are other potential extensions I have not yet thought of. The only way to reach any conclusions on the matter would be to try it.

C O N C L U S I O N S

CONCLUSIONS

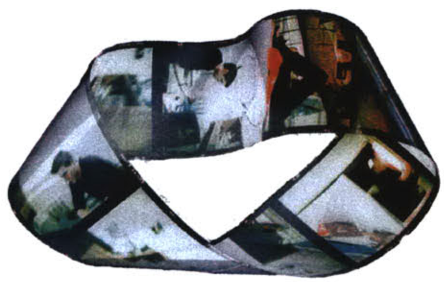
This thesis sought to create an interactive environment that allows users to both see and interact with the past. The presented past was not to be a complete history, but rather an elastic memory that represents more "interesting" events with greater temporal resolution than those of greater monotony. Despite the limitations highlighted in the previous chapter, the project did succeed in certain goals set out for it. The basic structure for the environment—the useful metadata, active capture and reactive environment—were all successfully appropriated for the implementation of this project and proved an effective model for a beginners attempt at the creation of this unique Simulacrum and serve as a proof of concept.

We no longer think of the history of cinema as a linear march toward only one possible language, or as a progressioin toward 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.

-- Lev Manovich
"What Is Digital Cinema? "

As highlighted in the Background section of this document the image, and more evidently so, the moving image has the ability to construct unique temporal and spatial relationships between the viewer and the representation. New Media theorists are continually saying that new visual technologies are changing the structure of these relationships, which were highlighted in the Introduction. This project demonstrates the use of these technologies—specifically sensor data, continuous capture, and real-time presentation—in combination with other technologies—segmentation, image composition, and data-basing techniques—to create a new such relationship between both individuals and a social community and their environment and past that was previously not possible. The metadata and active capture were used to create new forms of image categorization, create a new interface to image collections and add the human to the loop of these algorithmic processes. The segmentation, image composition, and data-basing techniques were used in conjunction with this to foster a social and temporal connectedness.

So while it's true that the introduction of these new technologies associated with the digital image are redefining codes of interaction, they are also enabling us to establish new and unforeseen experiences that create new codes of interaction, this work being just one example.





B I B L I O G R A P H Y

BIBLIOGRAPHY

Agamanolis, Stefan. *Isis Cabbage, and Viper: New Tools and Strategies for Designing Responsive Media*. MIT, PhD Thesis, June 2001.

Agamanolis, S., Westner, A., Bove Jr. V.M., "Reflection of Presence: Toward an More Natural Responsive Telecollaboration," SPIE Multimedia Networks. 1997.

Agamanolis, S., Piper, B. *Palimpsest*. http://www.mle.ie/hc/projects/palimpsest/, 2002.

Agamanolis, S., Cullinan, C. "Reflexion: a responsive virtual mirror," UIST Symposium. Paris, 2002.

Barthes, Ronald. *Camera Lucida: Reflections on Photography*. Noonday Press, 1992.

Baudrillard, Jean. "Simulacra and Simulations," *Jean Baudrillard: Selected Writings*, Ed. Mark Poster. Cambridge Press, 1988, pp. 166-184.

Benjamin, Walter. "The Work of Art in the Age of Mechanical Reproduction," *Illuminations*. Pimlico, 1999, pp. 211-244.

Bergson, Henri. *Matter and Memory*. Zone Books, 1990.

Borges, Jorge. *Labyrinths*. New Directions Publishing Corporation, 1988.

Bove, Jr., V.M. *Object-Based Media*. http://web.media.mit.edu/~vmb/obmg.html, 2003.

Buck-Morss, Susan. *The Dialectics of Seeing: Walter Benjamin and the Arcades Project*. MIT Press, 1991.

Butler, D., Sridharan, S., Bove, Jr., V.M. "Real-Time Adaptive Background Segmentation," *Proceedings of ICASSP*, 2003.

Crary, Jonathan. "Spectacle, Attention, and Counter-Memory," *Guy Debord and the Situationists International*, Ed. Tom McDonough. MIT Press, 2002.

Davis, Marc. "Active Capture: Integrating Human-Computer Interaction and Computer Vision/Audition to Automate Media Capture." *IEEE Conference on Multimedia*, July 2003.

Debord, Guy. "Report on the Construction of Situations and on the Terms of Organization and Action of the International Situationist Tendency," *Guy Debord and the Situationists International*, Ed. Tom McDonough. MIT Press, 2002.

Debord, Guy. *The Society of the Spectacle*. Zone Books, 1995.

Deleuze, Gilles. *Bergsonism*. Zone Books, 1990.

Deleuze, Gilles. *Cinema 2: The Time-Image*. University of Minnesota, 1987.

Deleuze, Gilles. "Plato and the Simulacrum," *October*, No. 27. Winter, 1983, pp. 52–53.

Foucualt, Michel. *The Archaeology of Knowledge*. HarperCollins, 1976.

Foucault, Michel. *The Order of Things*. Vintage Books, 1994.

Ishii, Hiroshi. "Tangible Bits: User Interface Design towards Seamless Integration of Digital and Physical Worlds," *IPSJ Magazine*, Vol. 42, No. 3. March 2002, pp. 222-229.

Jaeckel, Monika. *The Delegation of Perception*. Interactive Media, 2002. http://www.delegate-perception.net/

Johnston, John. "Machinic Vision," *Critical Inquiry*, August 1999.

Karahalios, Karrie. *Telemurals*. http://web.media.mit.edu/~kkarahal/thesis/projects/telemurals.html, 2002.

Kuhn, Thomas. *The Structure of Scientific Revolutions*. University of Chicago Press, 1996.

Kuramoto, M., Masaki, T., Kitamura, Y., Kishino, F. "Video Database Retrieval Based on Gestures and Its Applications," *IEEE Transactions on Multimedia*, Vol. 4 No. 4, December 2002.

Levin, T.Y., Frohne, U., Weibel, P. *CTRL[SPACE]: Rhetorics of Surveillance from Bentham to Big Brother*. MIT Press, 2002.

Lyon, Chris. *Encouraging Innovation by Engineering the Learning*
Curve. MIT, MEng Thesis, June 2003.

MacKay, W. E. "Media Spaces: Environments for multimedia interaction," *Computer-Supported Software Work, Trends in Software Series*. Wiley and Sons, 1999, pp. 55-82.

Mitchell, William. *The Reconfigured Eye: Visual Truth in the Post-Photographic Era*. MIT Press, 1992.

Ozer, I.B., Wolf, W. "Video Analysis for Smart Rooms," SPIE ITCom, 2001.

Manovich, Lev. *The Language of New Media*. MIT Press, 2001.

Manovich, Lev. "'Metadataing' the Image," http://

manovich.net/, 2002.

Manovich, Lev. "The Poetics of Augmented Space: Learning from Prada," http://manovich.net/, 2002.

Martinez, José M. "MPEG-7 Overview," International Organization for Standardization, http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm, July 2002.

Poster, Mark. «Foucault and Databases: Participatory Surveillance.» *Modes of Information: Poststructuralism and Social Context*. University of Chicago Press, 1990.

Sawhney, N., Wheeler, S., Schmandt, C. "Aware Community Portals: Shared Information Appliances for Transition Spaces," Proceedings of CHI. The Hague, Netherlands, 2000.

Sawhney, Nitin. "Situated Awareness Spaces: Supporting Social Awareness in Everyday Life," MIT, PhD Generals Examination, 2000.

Schwartz, H. *Media Art History*. Prestel and the Media Museum, ZKM, 1997.

Shaw, Jeffrey. *The Golden Calf*. http://www.jeffrey-shaw.net/, 2003.

Venturi, Robert. *Iconography and Electronics Upon a Generic Architecture: A View from the Drafting Room*. MIT Press, 1998.

Virilio, Paul. *The Vision Machine*. Indiana University Press, 1994.

all wrongs reversed.