Sonic Images and Visual Poetics

Exploring a methodology for sound and image relationships

by

Martha W. Swetzoff

BA, Social Anthropology
Harvard College, Cambridge, Massachusetts
1984

Submitted to the Media Arts and Sciences Section
in partial fulfillment of the degree of
Master of Science in Visual Studies
at the
Massachusetts Institute of Technology
June 1989

Signature of Author ______________________________
Martha W. Swetzoff
Media Arts and Sciences Section
February, 14, 1989

Certified by ______________________________
Glorianna Davenport
Assistant Professor of Media Technology
Thesis Supervisor

Accepted by ______________________________
Stephen A Benton
Chairman
Departmental Committee on Graduate Students

DISCLAIMER NOTICE

The accompanying media item for this thesis is available in the MIT Libraries or Institute Archives.

Thank you.

Sonic Images and Visual Poetics

by
Martha W. Swetzoff

Submitted to the Media Arts and Sciences Section
on February 14, 1989 in partial fulfillment
of the requirements for the degree of
Master of Science in Visual Studies.

Abstract

This thesis emerged in parallel from specific ideas within musical and cinematic traditions. In the course of developing a methodology for working with sounds and images I completed a half-hour video and sound piece, entitled "Tales of Love and Glory" (TLG), included as part of this thesis. This project brought up artistic dilemmas for which the existing models of sound production in the traditions of film and video failed to present satisfactory aesthetic and technical solutions. In the search for a new paradigm I explored the growing field of computer music. There, research and methodology encompass a level of scientific description of sonic materials, ideas about higher-level cognitive processes of auditory perception, and musical aesthetics. This approach offers a powerful set of theoretical and technical tools for artists, composers and sound designers engaged in creating a language of sound and picture relationships for motion pictures.

Included in this thesis is a discussion of two digital audio workstations: the AudioFrame, designed by the WaveFrame corporation, and the Synclavier, designed by the New England Digital corporation.

Thesis Supervisor: Glorianna Davenport
Title: Assistant Professor of Media Technology

ACKNOWLEDGEMENTS

The author would like to thank the following individuals and corporations:

Steve Cunningham
Steve Krampf
Andrew Currie
The WaveFrame Corporation
Ted Pine
The New England Digital Corporation
Curtis Roads
Adrian Freed
Tod Machover
Glorianna Davenport
Alan Lasky
Sabrina Birner
Bruce Middleton-Baird
Alan Ruttenberg
Rainer Boesch
Jean-Claude Risset
Gwen Gordon
David Heeger

Contents

PARALLEL STREAMS

Introduction

This thesis emerged in parallel from specific ideas within musical and cinematic traditions. In the course of developing a methodology for working with sounds and images I completed a half-hour video and sound piece, entitled "Tales of Love and Glory" (TLG). This project brought up artistic dilemmas for which the existing models of sound production in the traditions of film and video failed to present satisfactory aesthetic and technical solutions. In the search for a new paradigm I explored the growing field of computer music. There, research and methodology encompass a level of scientific description of sonic materials, ideas about higher-level cognitive processes of auditory perception, and musical aesthetics. This approach offers a powerful set of theoretical and technical tools for artists, composers and sound designers engaged in creating a language of sound and picture relationships for motion pictures.

I became interested in the language of sonic images in contemporary music. This direction of compositional thought embraces timbres, or sound colors, and sound-objects, as opposed to traditional musical notation, as the building blocks of composition. This branch of musical thinking has been greatly affected by the introduction of the computer as a means to analyze, order, synthesize or modify sonic materials. Works such as *Sud*, by Jean-Claude Risset, and *Mortuos Plango, Vivos Voco,* by Jonathan Harvey, excited me because they contained extremely supple transformations of naturally-occurring sounds. I became interested in the way, due to the precise control offered by the computer, perceptual illusions might evolve from such treatments. I became particularly intrigued with combining this approach to sound and music composition with moving pictures.

The first half of this thesis begins with a short essay on time. I then discuss the artistic conditions which provoked this search, the background of traditional sound use in motion pictures, and the work of the composers Trevor Wishart, Jonathan Harvey, and Jean-Claude Risset as examples of a new methodology.

The first section ends with a discussion of visual concepts that have informed this thesis work. These include, but are not restricted to, certain areas of photography and cinema. "Tales of Love and Glory" incorporates a diverse set of materials, and represents, a synthesis of many kinds of ideas. These are not simply defined by the specific format of the final piece, which could be described as;" a videotape, shot in film, with a musical soundtrack." Important sources of influence on this work also include painting, poetry, perception and psychology. I have tried, through this piece, to layer many kinds of meaning, in a way that is only possible through a broad, inquisitive approach to the sights, sounds and thoughts of this world.

As Bergman says:
I'm a radar set. I pick up one thing or another and reflect it back in mirrored form, all jumbled up with memories, dreams and ideas. A longing and a will to form.[1]

In the second half of this thesis I present a detailed discussion of the tools used for the production of TLG: two digital audio workstations. One of these is a new design, the AudioFrame, made by the WaveFrame Corporation, and the other is the Synclavier, made by the New England Digital Corporation. Here, I outline the history of integrated digital audio workstations, as well, and discuss interface and design issues that concern the test systems and future directions for tools in the context of post-production for film and video.

A SHORT ESSAY ON TIME

I think that what a person normally goes to the cinema for is time : for time lost or spent or not yet had. He goes there for living experience; for cinema, like no other art, widens, enhances and concentrates a person's experience--and not only enhances it but makes it longer, significantly longer.[3]

---

1 Bergman on Bergman, Bjorkman, Manns, and Sima, eds.
3 Tarkovsky, Sculpting in Time.

How can we enter into a discussion of music and moving pictures without considering their outstanding shared feature? Both unfold in time, take their form from it. What other human endeavors can so readily embrace serial and parallel constructions? It is technology that has allowed us to fix and manipulate time. Is the history of art, and perhaps of science, a road map of our struggle to escape its limits?

The fact of time is a burden we all share. As the one feature of our world which eludes our absolute control we are drawn to those processes that offer us some kind of refuge, however brief, from our powerlessness, and an escape from time's relentless attentions. Perhaps if we can conquer time, we can conquer fragmentation.

> *Time separates me from myself, from what I have been, from what I wish to be, from what I wish to do, from things, and from others.*[4]

A spectator surrenders a bit of his life to the experience of a film or piece of music. Media that have their basis in a static object, such as a painting, or a fixed moment, such as a photograph, cannot make the same demand. If a work unfolds in time, how do we distinguish it from other aspects of our life, which also occur in real time? Time passes while we watch a movie, or listen to music. Is it the same kind of time which passes while we wait for the bus, do the laundry, read a book? Tarkovsky's quote suggests that films offer a form of vicarious experience, a way of literally gaining more 'time' in life by witnessing lives recapitulated in finite, distilled presentations.

For many people watching the world on the screen, be it in the cinema or on the tv, remains their broadest interaction with life and emotion. How often does one feel that one has actually had a certain experience because of seeing it on tv or in a movie? Films enrich our experience of the world. But how? Films--and here I am not distinguishing between video and film--appeal to us because they fix time, and because they offer us a privileged view, since time transpires in the filmmaker's terms. The filmmaker

---

4 Sartre, Being and Nothingness

modifies time through pacing, in the editorial process, and choices about camera speed and angle made in shooting the film. We can return to examine what has been fixed in moving time, in a way we never can return to our own experiences, or memories. Through montage, or layering techniques, a motion picture can compress or expand time. And films present a view of the world which, however close, is not our own--it is that of the filmmaker. Thus film can enrich our experience both by the distinct, selective nature of what it shows us, but also by the addition of a point of view which expands our own.

Could we perceive a sound if it had no duration? We understand sound to possess characteristics which vary over time. These time-varying properties have a shape known as an *envelope*. For example, the amplitude envelope is often partitioned into three stages: the attack (onset), sustain, and decay. The changes of frequency and amplitude over time define the sonic world. Though we sometimes hear music in snippets, we do not hear in single frames. Music presents many different sound 'snapshots' to a listener, often in quite repetitive structures. Marvin Minsky asks, "Why isn't music boring?"

> *Is hearing so like seeing that we need a hundred glances to build each musical image?*
>
> *...Then why do we tolerate music's relentless rhythmic pulse or other repetitive architectural features? There is no one answer, for we hear in different ways, on different scales. Some of those ways portray the spans of time directly, but others speak of musical things, in worlds where time folds over on itself.* [5]

He suggests that music provides a way to learn about time. We tolerate repetition in music because our mind's agents concentrate on matching patterns and discerning variations. The exploration of how parts fit or vary represents an important activity in our apprehension of the world. Since much of musical activity involves building larger structures out of smaller

---

5 Minsky, "Music, Mind and Meaning", in C. Roads, ed., The Music Machine

ones, he further observes that perhaps this "drive to build mental music structures is the same one that makes us try to understand the world."

Conjectures about the mechanisms and goals of our mental processes in the processing of sound and visual data enter treacherous territory. However, developments within the field of cognition offer artists working with time-based media an important resource for insight into the nature of their activities. The kinds of questions asked in this level of inquiry, such as "Is hearing like seeing?" "How does the mind treat objects of hearing in relation to objects of sight?", may not have concrete answers yet. These questions are intriguing and stimulating not only scientifically, but artistically. As research in the area of cognition and human perception advances, artists can enter a deeper understanding of their own creative impulses and how they might communicate them most effectively to others.

Questions about hearing and seeing hover around a work like "Tales of Love and Glory". In the course of this work I found myself wishing I could find some kind of profound principle about perception, which could imbed the work more deeply in the realm of memory and dreams. Ultimately, I reached for the tools I knew about in music and filmmaking to suggest the sensory world I was after, but the curiosity remains.

I am intrigued by the struggle waged by human beings to accept the fact of time. We have tools to trap time, but time also traps us. Our perceptual apparatus seizes time in discrete sensory units, and re-delivers these bits seemingly at random, through the experience we call memory. Sometimes our point of view of the world can be frozen in time--suspended around the point of view we had during a moment or period of great intensity. Miller (1981) notes that children who undergo loss, separations, parental traumas or abuse, often form a bond with the natural world as a source of stability and consistency which does not threaten delicate parental relationships. She notes that they might be able to " remember exactly how they discovered the sunlight in the bright grass at the age of four, yet at eight be unable to 'notice anything'...."

"Tales of Love and Glory"

"Tales of Love and Glory" is a half-hour videotape, with an experimental soundtrack. The tape was shot initially in super-8 film, in color and black-and-white, then transferred to videotape. I collected the image and sound materials over a period of two and a half years.

The aim of the piece was to explore childhood and adulthood through the perceptual phenomenon that Miller alludes to above--that as children our perceptual frame can be so strongly influenced by events which we can barely remember. Our struggle with time as human beings arises out of our awareness of past and present and the emotions evoked by the passage of time. I wished the work to suggest, not describe, the state of tracing back the threads of memory through the sensations of the present. Although the title of the work suggests a narrative format, it refers rather to a desire to invoke many stories, the stories we all possess as human beings with personal histories.

This piece follows after two other experimental short pieces intended to explore certain issues of sound and image relationships. One, "AutoPainting", was an experiment in combining processed and unprocessed images, music and spoken text. In many ways, this piece was a sketch for "Tales" . The images were processed with a real-time video effects processor called a Fairlight Computer Musical Instrument. The un-processed, source images originated in documentary and staged situations, and thus also formed a kind of exploration in narrative and cinema verite techniques. The music was executed using MIDI and traditional analog audio studio techniques. The MIDI score was generated using the Performer program, and synchronized with a Roland SBX 80 SMPTE-to MIDI interface.

The second experiment was a video piece, "Grace" (6, minutes, 1987), which used only the processed video materials. This piece explored purely abstract forms and organization in the picture domain, and used FM timbres designed by me and Thomas Trobaugh of the Music and Cognition Group at the Media Lab. Tom and I collaborated on this piece. The soundtrack here was done purely as a MIDI score to video using the same

SMPTE-to-MIDI setup--there was no spoken text or voice. Our plan was to try to compose with the abstract visual forms rendered by the Fairlight and mirror this process in the timbre design and composition. We designed the timbres according to various qualities of color, abstract shapes, and motion that we felt typified each set of images.

"Tales of Love and Glory" represented a synthesis of the technical and aesthetic ideas I had developed from these two previous pieces. I gathered the materials from many locations, including New York City, Maine, New Hampshire, Cambridge and Georgia. The sound elements originated either as wild sound from shooting locations, as sound effects which I collected very specifically for the composition, or from optical sound effects libraries. There are several categories of images: documentary, "observed" reality (fig 1-6); personal, subjective, "diary" (fig 7-18); and staged (fig.19-30).

The sources of artistic dilemma

This project initiated in the following thoughts; How can I make a film which talks about time but doesn't contain any words? Could I imbed the text of the piece in the soundtrack through the timbre selection and acoustic relationships in the composition? To what degree can sound convey meaning on its own, without the spoken word, and without the traditional notion of the narrative? Is there another way to "tell a story" with moving images and sounds which depends less on literary tradition, or the motivic use of music, and instead arises from inherent properties of sound structure and visual stimulation?

TLG is far more ambitious in scope and scale than my previous projects, and required a change in method based on that alone. I realized as well that this project, due to the extent to which the soundtrack would carry the weight of the piece, represented an opportunity to integrate further various ideas that had arisen through the study of developments in computer music. I became very interested in the idea of an integrated digital audio workstation as a tool to allow very precise manipulations of both sonic materials and the sequencing of these materials in the video environment. I knew at that time that I wished to combine many sorts of sounds in the soundtrack, and that I wanted the soundtrack to have a density and complexity that would all be invented in post-production. I wished the soundtrack to combine synthesized elements, ambience and processed natural sounds into a fundamentally musical form, but one which could slip in and out of seemingly synchronized sound. I hoped to use a great deal of human voice manipulation, but have the voices remain on the level of utterances, and to avoid a direct voice-over treatment of text if I could help it. I realized that I would need a powerful tool, or set of tools, in order to subtly achieve these goals. I would also need a new way of thinking about sonic materials.

As an example of the matching of visual and auditory material I will describe the opening section of the work. There are five sections in total. This opening section has a text inspiration, which shapes it structurally in visual terms. This text is an excerpt from a short story called, "Lyompa", by a Ukrainian writer, Yuri Olesha. The story describes the final hours of an ill man, close to death, who realizes that the things which had surrounded him in his active life now exist only as abstraction in his mind. He possesses only names. He exists outside of the laws which govern objects, such as gravity and causality. He realizes that he had gone through life feeling that his "eye and his ear ruled things". "I thought the world would cease to exist when I ceased to exist. But I still exist. So why don't the things?" He contrasts his plight with that of a child, who wanders into his room. The child has no names for anything. The world, in its entirety of sensations, rushes to meet him.

This text itself does not appear in the context of the sequence. But the ideas are present visually. The first shot (in a black and white sequence) shows a close-up of a photograph which is revealed to be part of still life on a table. Through a series of dissolves, the other objects disappear form the surface of the table, finally leaving a bare white surface. The other objects include a small glass of milk, a bowl of steaming soup, a lamp, a toy miniature rocking horse and cradle. The latter is filled with wooden scrabble letters, but only the letters l, o, v, and e. The black and white sequence fades out from the bare table and cuts into a color shot taken from behind the head of small child as she is being pushed in a stroller through a park. The shot has been undercranked--shot at nine frames per second, so the effect is somewhat jerky and pixillated, and results in a feeling of speed. Focus shifts with this technique, due to slightly unsteady handling (I was positioned right behind her head, with the camera below my waist, viewfinder closed, at the widest angle setting.) The child's head is centered in the frame, and her looking about thus presents the primary action. As she passes people and things, they glance at her or smear by. I occasionally cut to versions of this shot which I slowed down in video. This adds a certain feeling of suspension to distinctive moments, such as when another child walks by and looks at her. The sequence fades out from one of these 'distinctive moments'.

Bits of the Olesha text appear throughout the work as sampled and manipulated voice fragments. Here, they appear as samples from a reading, which were sampled running backwards, through a digital delay. They are played back forwards, and have been looped and pitch-shifted. A chorus effect has also been added and stereo-panned, to accentuate movement of the sibilant parts of speech. These whispery, voice forms, which are recognizable as voice although not articulating discernable language, dominate in the black and white section. The shifting sibilants synchronize fairly closely with the visual dissolves. A gong-like timbre also comes into this section in synchronization with visual events. By the cut to color, the voice has become a more subdued element. Another timbre rises to audible surface, which has been built from a sample of rubbing a crystal glass, with the attack removed, in combination with a synthetic timbre produced through frequency modulation. A wind effect appears here, produced from an optically-stored sound effect, to which vibrato and a modulated, ramped attack were added. A slowly-arpeggiated vibraphone-like timbre, produced through FM synthesis, appears here, also. It appears in different transpositions, and in different modifications of partials. One has been delayed and stereo panned, for instance, another delayed and detuned, etc. In its higher register it evoked tinkling bells. A sampled clarinet, again with attack modifications and slightly chroused, also adds a melodic color to this section.

I should add that some of the voice fragments that appear in the work as sampled instruments were first processed with a system developed by the

Swiss composer Rainer Boesch. This system consists of a custom DSP card which fits into an Apple II computer, and performs manipulations on incoming audio real-time. I used this to process readings of the Olesha text, other texts, and audio from several video-8 interviews I shot with some of the individuals who appear in the piece. This additional step of audio manipulation allowed me to significantly extend the vocabulary of the voice-based materials.

The Fragmentary Behavior of Memory

"Tales of Love and Glory" concerns itself with the fragmentary behavior of memory. Considerations of "time" anchor the piece. (as both the subject and the mode of delivery.) I create a serial construction of images that are linked, end to end. This structure sits on top of a series of stacked and linked sound objects. All of these move in time together. The images will be understood in certain kinds of ways, having to do with changes that occur over time, and the relationships of shot content--understandings having to do with the conventions of cinema. The sound will be understood in yet another series of ways which are to me more subtle, but that clearly involve changes and transformations over time as well and invoke another kind of perceptual activity. In combination with the visual stream the soundtrack can promote the kind of stimulation of present and past sensations that constitutes, in my work, the basis for the experience. This notion of flipping tenses between past and present, between observed and subjective realities (what is seen and what is felt) constitutes the crucial dialogue within "Tales of Love and Glory".

> *Time is the substance of which I am made. Time is a river that bears me away, but I am the river; it is a tiger that mangles me, but I am the tiger; it is a fire that consumes me, but I am the fire.*[1]

What is moving about that passage? The tautology of our being both subject and object, of time folding in on itself. Is this a passage about musical things?

---

1 Borges, "A New Refutation of Time". In A Personal Anthology, pg. 64.

This piece ultimately represents a paradigm change, in the sense that Thomas Kuhn describes as being the prerequisite for scientific revolution, when," an existing paradigm has ceased to function adequately in the exploration of nature to which that paradigm itself had previously led the way[1]." I felt that the the familiar methods of film and video sound production could not function as a paradigm in the production of this piece in either the aesthetic or the technical sense. I needed a new methodology. This by definition encompasses theoretical and practical concerns, and the following describes my search for a new set of creative principles and the tools with which to execute them.

Musical streams

I have earlier described some of the procedures used in producing the soundtrack. My concepts for the construction of the soundtrack did not arise from the tradition of composing for motion pictures, but from some of the ideas developed through the evolving discipline of computer music. I did carry into this work vestiges of my prior experience in film sound editing, which influenced me in determinations about sonic materials, placement of certain kinds of sounds against picture, and ultimately in my choice of tools to use for this project. The experience of hearing a musical work where the essential nature of the sonic material constantly yet subtly shifts is one of great ambiguity [2]. An atmosphere of compelling suggestiveness arises out of a situation where the mind is not really sure what it is listening to, *exactly*.

In the context of applying the dimension of sound to motion pictures, the power of an ambiguity independent of the objects shown on the screen seems tenuously explored in relation to musical compositions such as *Sud,* by Jean-Claude Risset and *Mortuos Plango, Vivos Voco*, by Jonathan Harvey. As my understanding of some of the sophisticated techniques in the computer music domain grew I realized how primitive the tools were for manipulating sound in the environment of motion pictures.

---

1 Kuhn, The Structure of Scientific Revolutions, pg. 92.

2 Harvey, "The Mirror of Ambiguity", in S. Emmerson, ed. The Language of Electroacoustic Music

Traditional film sound technique

As the sound film evolved, it came to rely on a literal soundtrack--one where sound illustrated the actions on the screen in order to complete the illusion of reality. Synchronized sound supports the realism of the cinematic experience. Music exists traditionally as a sort of emotional instruction to the audience about how to feel throughout the film. The literal approach is starting to change, notably in the work of sound designers such as Alan Splet, who designs the soundtracks for David Lynch, and Walter Murch, who has designed several soundtracks for Lucas and Coppola. Murch designed the tracks for *THX 1138, The Conversation,* and *Apocalypse Now* , among other films.

The narrative film as we know it in this country represents an extreme of sorts in this regard, as it suffers under more rigorous concerns of market place than experimental cinema, video art, or the European narrative filmmaking tradition. Foreign filmmakers such as Godard, Bresson and Wenders display a consistently open attitude to the use of sound in their filmmaking. Filmmakers on the edge of the form, such as Peter Kubelka and Dziga Vertov, present perhaps the purest examples of pushing sound in the film idiom, as we shall see further on in our discussion. Splet and Murch have had access to the state of art in film sound production, due to the economic scope of the projects they have been involved with. Crucial to the freedom which they have experienced remains the aesthetic concerns of Lynch, Coppola and Lucas, who in their different ways have supported the extension of sound beyond the traditional Hollywood mold. But even on the scale of Murch and Splet, the technology for manipulating sound in this context is still predominantly imbedded in magnetic media, which presents an analogous set of conditions to musique concrete.

Musique Concrete

Musique concrete is a technique of sound composition initiated in 1948 by Pierre Schaeffer. It was made possible by the technology of disc recording, followed soon after by audio tape recording. These technologies allowed sounds to be fixed, and played back repeatedly. Schaeffer cut and re-combined recorded materials with the explicit intent of exploring this process musically. The editing was not directed toward the illustration of a world exterior to the sonic one itself, nor were the sounds chosen and organized to provide background effects to a spoken scenario. Since there did not exist at that time the plethora of external effects processors that surround us today, Schaeffer's vocabulary originated in a language of editing and layering. Here, timbre changes could occur through changing the speed and direction of playback. But he also used innovative recording techniques, specifically by miking sounds extremely closely, thereby amplifying many "microscopic" sounds. Timbre changes can also occur through miking techniques, by recording off the axis of the microphone. The use of microphone proximity and directionality thus formed another aspect of the possible timbre palette available to the early practice of musique concrete.

Schaeffer established and explored the basic techniques of electronic manipulation, and demonstrated the potential of electronic apparatus for expanding the sonic world of the composer[1]. In this sense, his activity set the stage for electronic synthesis, which took its first steps in the early 1950s. He also founded GRM, Groupe de Recherches Musicales, a facility for sonic research which remains active today.

Schaeffer developed two other ideas which came to have an influence beyond the methodology of music concrete. One articulated the *sonic object* as a compositional building block. We can understand this term in several ways. One is the extraction of specific sound fragments from the recording and their use as instruments. But this is also a more general concept. It stands as a unit of organization for a musical world that is beyond notes, such as a *sound mass,* or *sound clouds.* He also realized that a simple operation, the removal of attack characteristics, could have a profound effect on the nature of the sound and how it was perceived. In this manner, bells could become organs. Or a distinct sound, such as rubbing the rim of a crystal glass and producing a tone, will be transformed into an ambiguous sound (just the tone, how is it produced?) that contains elements of the familiar without any of the expected cues. By embracing sounds from the

---

[1] see Griffiths, <u>Modern Music</u>, pg 31

world--'observed' sounds--musique concrete offered the world, in all its acoustic diversity, as a source of musical exploration.

Musique concrete created a flurry of interest amongst the post-war musical avant-garde. Its impact hit first in Europe, where many composers, notably Boulez and Stockhausen, were trying to extend, and re-invent, fundamental musical notions. Interest in serialism was quite strong at that time. Schaeffer's experiments were broadcast on French radio, and contributed to the general ferment. In an atmosphere already cloudy with experiments in rhythm and timbre, which were nonetheless confined to the possible repertoires of traditional instrumental technique, Schaeffer's inventions presented an almost pure manifestation of these ideas. The attraction to his method drew many composers to Schaeffer's studio. Messaien brought several of his pupils there, and himself produced a 15-minute piece, *Timbres-durees.* Boulez and Stockhausen also produced some short pieces there. This seemed to be a mutually productive interchange, for Schaeffer's vision of future directions grew. In an issue of *La Revue Musicale,*[1] he postulates some specific directions which seem quite prescient. One considered the technical ability to focus a very complex sonic event into a very simple one--at that time still an imagining. The other concerned the possibility of "the treatment of duration as a variable with the same capacity for complex relationships as pitch[2]."

Musique Concrete in the cinema

The influence of the ideas of musique concrete infiltrated into cinematic activities as well as continuing a part of musical thinking. The notion of the sound object, and a related consideration, of a sonic landscape have been extensively examined by Trevor Wishart[3]. In the consideration of a new language of sounds and moving pictures it is worthwhile to consider how Wishart and Murch differ in their considerations of the legacy left by Schaeffer and the activities of the Groupe des Recherches Musicales.

---

[1]Griffiths, p.72

[2]Ibid., p.77

[3]Wishart, On Sonic Art and Wishart, "Sound Symbols and Landscapes," In Emmerson, ed., The Language of Electroacoustic Music

I was very interested in the early fifties by Pierre Henri (Schaeffer) and those guys in France. I had copies of their records and was consciously working in that style.[1]

Walter Murch came across his first tape recorder at the age of eleven. He played with it throughout his teens. It did not seem possible, however, to make a living out of this interest. He ended up at film school at the University of Southern California, in the same class as George Lucas. Murch's approach to the construction of soundtrack represents a logical extension of musique concrete methodology into filmaking. He has developed techniques to create sonic objects, manipulating them in space, or, using techniques of creative re-recording to suggest a spatial definition according to the amount of presence he believes a sound ought to carry at each moment in the film. He gathers the majority of his sound material himself. His work also involves the concrete sounds of the world and transforming them into sensory experiences.

In his notes for THX 1138: *...think of musique concrete. Do not be limited by what you see on the screen; it would be better if nothing were related than if everything were.*[2]

The decisions made by a sound effects editor, or a sound designer, all at some point must concern themselves with two major issues, truth and drama. The design of the effects must be truthful in two respects-- emotionally, and perceptually. For instance, in a classic example, the actual sound of a gunshot is not that compelling. It may sound like a car back-firing, or an indeterminate popping. But this sound has a color which comes from our knowledge of guns. The sound emanates from a machine that kills. That is another kind of truth about that sound, we could call it an emotional truth. Sounds exist in our minds as well as in the world; the sound editor must guess the expectations of the audience.

---

[1] Rockwell, All-American Music, p. 155
[2] Ibid., p. 157

This becomes a challenge because sounds do not necessarily record the way we hear them in the world, and thus can lack credibility when placed to picture (even, paradoxically, if they were recorded on location). The art exists in how closely to match the sound as we imagine it, and expect it to sound, both from the reality constructed by the picture and from the dramatic reality that the film aims to construct. The sound editor makes decisions about how near or far to position the expectation of a sound which describes an action. Choices are shaped by the a balance of the two concerns; is this a moment when we need to establish something about the physical space which surrounds these characters? Or is this moment important dramatically to introduce an emotional element?

Sounds that don't match seem wrong because they carry associations that contradict the screen reality. If the door slam doesn't fit the expectation of how that door should sound, questions suddenly arise. Why does that door sound like a heavier door, closing in a reverberant room? It is shown to be a hollow core door, closing in a small bedroom...is it supposed to be that way? Does it mean something? We could perhaps regard this endeavor of the sound editor as: what questions do you want viewers to have in their minds at this moment, while looking at x? In sound effects work, the final judgements about sound design remain ultimately formulaic and somewhat alchemical. That is, not based in a scientific understanding of the structure of sound but in a kind of tradition of experts, who have built up their expertise over years of practice at determining what sounds successfully pull an emotional response out an audience.

Murch says, "The sound is what's in your mind, not what is on one piece of magnetic tape[1]." The sound designer traffics in ultra-reality. Murch talks about giving "the impression of that density of reality". How might that reality change if one were to approach it from a musical standpoint? What would happen if the tools available allowed a finer level of access--beyond strips of tape with discrete sounds recorded on them?

[1]Fox, "Walter Murch --Making Beaches Out of Grains of Sand,", Cinefex

Two themes, which can be seen as having an origin in the domain of musique concrete, have maintained a durability throughout changes in technology and method. Both of these have been important to my work, and, I believe offer a stimulating jumping off point for a discussion about expanding the dimensionality of the sound domain in motion picture production. The first of these, the notion of the sound-object, initially formulated by Schaeffer, remains a foundation for a certain kind of musical thinking concerning materials and a vocabulary of transformational process on them. The second of these consists of the notion of a sonic landscape.

Wishart and Aural Landscape

As a composer who uses such fragments of the real world as birdsongs and vocal utterances, Wishart has developed a detailed contemporary exegesis on the use of what he terms 'representational' sounds. His approach represents another methodology which illuminates some issues in the treatment of audio materials.

Wishart points out that we are inclined, when hearing a sound, to associate it with a source. Schaeffer's original notion of the 'sound-object' placed great importance on the separation of the object from its context, whether that be its source or its cause. Wishart suggests that sounds exist in a landscape, and we might define that as "the source form which we imagine the sounds to come". The loudspeaker has allowed us to construct virtual acoustic space, "into which we may project an image of any real existing acoustic space, and the existence of this space gives us new creative possibilities."

In this conception, the loudspeaker mediates the audio material and allows the creation of aural landscape which exists exclusively in the space bounded by the speaker. It is not a space that has to adhere to any external laws. Here, of course we see already a distinction in context from the tasks of the film sound designer as described above. The latter does have an external constraint which he must acknowledge in some way. But the overall point of Wishart's, which is that this landscape can maintain aural images as source of metaphoric or perceptual discourse via a plethora of techniques offers a rich source of ideas for the sound designer. Acceptance of these ideas does require that the composer working with pictures release himself from the responsibility to articulate every sound, even the most dramatic sounds present in the image, and consider the creation of another landscape which surrounds the image. This landscape may resemble, or refer to, the landscape suggested by the image, but it may also only refer to in the most subtle manner, more to engender a state of looking than to convince of reality. In fact, playing with the notion of acoustic reality becomes a powerful extension of this approach.

Wishart outlines some distinctions which illustrate how this landscape can be constructed and then re-defined. As an example he describes a forest environment as an acoustic space defined in width by the placement of speakers and in depth by changes in amplitude and high-frequency components, and perhaps increasing reverberation. The sounds of various birds and animals could be introduced and the perception of an inhabited, and in some sense, real, landscape would become even stronger to a listener. Then, if the animal and bird sounds were replaced by arbitrary sonic objects, a slightly different landscape would emerge. He points out that since the disposition of the objects has not changed, the overall acoustic image of 'forest' remains. Yet, ultimately all the sound sources might not be in a any sense real. He describes this as an imaginary landscape of the type 'unreal objects/real space.'

The second imaginary landscape in this scheme could be one where the original sound objects are subjected to arbitrary changes in amplitude, reverberation or filtering. This would produce a landscape of the features 'real objects/unreal space'.

A third type is evoked through addressing the nature of the relationships between the sound-objects. For instance, let us suppose that the creatures inhabiting this landscape had some kind of reasonably correct ecological relationship to each other. Editing or mixing operations that resulted in song sparrows duetting with train whistles, or lions, would constitute an world where the acoustic space still remained constant, that of the forest, but the relationships have now become possible only as an invention.

Wishart calls this landscape surrealist, for just as this movement in the visual arts depended on the introduction of unrelated objects into a visual space, this use of sonic manipulation brings together "normally unrelated objects in the virtual space created by the loudspeakers..."

The construction of aural landscapes can be affected by many other factors. Spatialization can be a potent effect, and computers have added a great deal of control to extending this technique. Sounds which rotate in space around the listener suggest that the listener is a point at rest. Wishart has explored quadrophonic space in some depth and points out that rotating an entire frame of reference creates a contrary effect for the listener; that he or she is spinning in an opposite direction while the frame of reference remains still. Other ways of changing listener perspective include changes in miking distance, which not only communicate changes in the distance to the sound source, but also carry psychological or social distance.

Not all of these techniques are practical for adaptation to motion picture presentations. Spatialization remains an expensive and problematic feature to attach to a film or video work. This remains a luxury confined largely to major feature films, such as *Apocalypse Now* . Smaller theaters which might run an experimental work lack adequate facilities. In fact, even the large theater equipped with 6-track dolby have problems maintaining the proper setup of the system in the absence of well-trained staff. Tuning the theater becomes compounded with economic factors which make breaking even a problem even without additional technical adjustments.

In common sound usage in film the directionality usually remains frontal, focused towards the screen. Speakers are placed behind the screen to heighten the the notion of screen as sound source--to further unify the picture with the sound elements that theoretically belong there. Video suffers perhaps more horribly in this regard, because often speakers are mounted, or placed right next to the screen. The limited screen size of video can make a sonic presentation that is too lateral seem dislocated from the image--if this is not intended one has to minimize spatial effects in order to preserve a senses of unity.

But the notion in general of changing the viewer/listener's aural perspective is one that can be achieved in other ways than through spatial manipulations, although this is the most dramatic technique. The use of miking distance as a tool offers a huge extension of vocabulary in this regard. Our perception of aural space can be deeply affected by the proximity of the mike to the source in addition to the associated timbre changes discussed in relation to Schaeffer. Microphone design has evolved since his time to include a wide number of different polar diagrams, which have a resemblance to camera lenses and filters. These basically fall under the following divisions, although there are many others: cardioid patterns (directional); supercardioid (very directional); and omnidirectional (not directional). These microphone patterns emphasize different spectral characteristics of the recorded sound. This can be a very simple and effective technique to introduce the idea of aural space, without requiring expensive technology.

Sound images as metaphors

How can we *understand* sound? If we take the point of view of cognitive psychology, that " all mental activity is mediated by internal (or mental) representations"[1] we can begin to consider hearing as some form of mentally representing sounds. Such psychologists as Langer (1942)[2] suggest that symbolization occurs on the level of a basic human need, and that it forms a fundamental act of thought and of mind. In this view, the type of sensory input is not as important as what occurs in the mind to the products of that input. Symbols focus and structure belief. If we assume that sounds can function on the level of symbols, then the arrangement of sounds with this in mind can touch on a profound aspect of our nature as human beings.

Musical form, or patterns over time, and the sonic characteristics which belong to sound objects thus both potentially could become assimilated as symbols. This begins to suggest some interesting analogies to cinema

---

[1] McAdams "Music: A Science of the Mind? " pg. 16

[2] As quoted in Ibid.

regarding *montage*, the joining of units of picture elements or *shots*. Changing the context of sound objects, that is, their relation to each other and their relation to the landscape which surrounds them, can result in changes of aural perspective regarding meaning. Sound will become interpreted differently according to what surrounds them.

When creating music from representational sounds, this becomes particularly pronounced--a sense of content develops because material becomes contextualized by the act of arranging it in certain orders and patterns. It can easily result in a form of narrative interpretation. Wishart, for instance, points out that the sound of heavy breathing heard in the context of a musical instrument might suggest a performance of extended vocal technique; heard in a quiet ambience it might carry sexual connotations and heard in the context of body hits would seem to describe a fight. In each of these situations the listener occupies a different position in relation to the acoustic activity. Can this property become useful in expression?

In motion picture terms, adjacent images can suggest wildly different interpretations. The Soviet filmmaker, theorist and teacher, Lev Kuleshov, revealed this attribute of film montage while teaching a workshop not long after the 1917 revolution[1]. Raw film stock was so scarce that they used old footage for editing exercises. In one experiment, a convincing senses of unified narrative was created by intercutting material shot at different times and locations, resulting in what Kuleshov termed, "creative geography." In perhaps the most famous of these experiments, footage of the actor Moszhukin, consisting of three identical shots, was intercut with shots of a plate of soup, a woman in a coffin, and a little girl. Audiences reacted to the sequence by remarking on the strength and subtlety of the acting--*now he is obviously hungry, now sad, now feeling affection.* The central shots of the actor did not change, only the images surrounding these shots.

---

[1]Monaco, James. How to Read a Film, pg. 309

Music has its own kind of narrative tradition, where music stands for narrative action, called program music. Richard Wagner had an understanding of the importance of developing form over time in relation to the expression of a narrative idea. This is not unlike myth, or storytelling, where the style and duration of the delivery become integral to the content. Myths articulate abstract concepts and relationships through the use of familiar cultural entities. Thus animals, humans or even inanimate objects acquire certain characteristics through myth. Wagner, in his use of the leitmotif, links musical structures to people, ideas, or things, and thus creates a kind of mythic discourse within musical structure.

Technology has made it possible to import compelling sounds into musical activity that carry their referent with them--that is, the composer does not have to invent a musical parallel as Wagner did, in order to point to the object symbolized. An actual representation of the object can appear instead. This has vastly increased the creative possibilities for this direction of musical expression. The construction of a metaphorical acoustic space creates an arena for a different kind of listening than the pattern-matching model Minsky describes. (Is this what he means by "hearing on different scales?") The introduction of digital techniques, high-quality sampling, sound analysis and editing techniques enables sound artists to explore this far more subtly than through the use of tape techniques.

Moving from the tape recorder to the computer

Computer technology has allowed composers to explore the symbolic power of sounds through structure far more delicately than with the old tape splicing techniques of musique concrete. The evolving science of psychoacoustics presents an additional set of ideas from which to structure our understanding of sound. The switch from the tape recorder to the computer allows the endeavors of composing and psychoacoustic exploration to occur as mutually illuminating activities on a deeper level than before.

Analysis of concrete sounds and synthesis based on this sonic data can yield images of varying degrees of proximity to the original sonic image. The event of transformation can itself become the focus of composition and musical exploration. The power of this technique lies in the possibility of shaping meaning out of seemingly vague sound objects, which can evolve into distinct identities, then shift subtly into a new one. The sounds of the world can become musical building blocks of far more plastic and expressive nature than in tape recorder based manipulations because the modifications can happen on the level of structure.

The expressive possibilities of this approach are best illustrated by the work of two composers, Jonathan Harvey, in his piece, *Mortuos Plango, Vivos Voco,* and Jean-Claude Risset, in his piece, *Sud.*

Mortuos Plango was realized at IRCAM, and consists of two sets of sonic materials, the sound of the great tenor bell of Winchester Cathedral and the singing voice of a young boy, who sings the text engraved on the bell. Harvey wished to contrast the dead voice of the bell with the living voice of the boy. The spectrum of the bell was analyzed with a fast Fourier transform program at IRCAM which was part of an interactive sound analysis program called S, developed at Stanford.[1]The bell was analyzed 1/2 second after it was struck, and this profile became the structural basis for the piece. The spectral analysis rendered a sixteen harmonic partials, out of these Harvey chose eight pitches from which to develop eight sections of the piece. In addition to the frequencies present at a typical moment of striking, Harvey included another spectral characteristic which is heard as a secondary strike note, this partial at 347 Hz had a slow decay and a beating component, but as a secondary strike artifact was not present in the analysis. This was determined by hearing and matching the frequencies present in the analysis.

After the stage of analysis with the bell, Harvey proceeded with synthesis and mixing using a version of the programming language Music V. He first synthesized the bell spectrum, and then could perform operations such as altering the envelopes (this could "turn the bell inside out", by making low partials which normally decay slowly decay very quickly), and

---

1 see Harvey, J. " Mortuos Plango, Vivos Voco: a realization at IRCAM", Computer Music Journal, Vol.5, no. 4, winter, 1981.

transforming from one bell sound to another by making glissandi pivot through a center tone. He also used the computer to read back the sound file in different ways, either forwards of backwards, and with the option of continuously varying the speed. Rhythmic patterns and spatial movement could also be devised.

The boy's voice was treated in a number of ways to effect subtle transformations. Initial recordings were made of the boy singing and chanting the text and singing the bell partials as rendered in the analysis. Harvey also simulated these sounds using the CHANT program developed at IRCAM by Gerald Bennett and Xavier Rodet. This is a singing synthesis program. Another technique used was to digitize recordings of the boy singing vowels, loop the sound files, and add pitch and amplitude contours "analogous to those applied to the sinusoidal components of the bell spectra." The boy's synthetic voice thus sang on the bell's partials. If bell-like envelopes were then applied to these boy-singing-bell, yet another level of transformation could result.

The overall repertoire of this piece and the process of developing it is yet more complex than my encapsulation. The effect of this detailed labour is an almost hallucinatory alternation of the perception of "boy" and the perception of "bell". Very startling moments occur when these two sounds seem simultaneously present, a duality unified. These moments throw the mind into doubt about the objects of its perception, and the assumptions it makes about the identity of these sounds. The sound of the human voice is one which humans can discriminate on a very fine level--on both the level of timbre and structure.[1] Harvey arranges the interplay of timbres to capatize on this subtle level of discrimination, and the illusions produced here thus occur

The care with which Harvey arranged these materials according to the harmonic structure of the materials takes the process from a sound effects exercise to a musical experience. Here, the computer provided access to

[1]McAdams and Saariaho, "Qualities and Functions of Musical Timbre", in Proceedings of the International Computer Music Conference, 1985.

data as well as the means to make musical structures from this data. His method stresses the ambiguity of sonic materials in order to explore the limits of human consciousness and perception. It is a method with roots in a rigorous understanding of sonic and musical structure, and colored by Harvey's sensitivity to the spiritual dimension of human experience.

Sud presents another approach to extracting musical forms from natural sounds. Jean-Claude Risset has been active in the application of computers to musical composition, timbre analysis and synthesis, and psychoacoustics since 1964.[1] He was initially drawn to the computer to expand "the functional use of timbre in composition", and started off his research activities at Bell Laboratories in 1964, where he worked on synthesizing brass timbres. His discovery that these timbres are characterized not be defined spectral characteristics but by specific changes in spectral distribution over time enabled the first synthesis of these otherwise elusive timbres. He has been particularly interested in auditory perceptual illusions, such as the Shepard tone, as well.

Sud is built up from a small pool of original sounds, some natural sounds recorded near Marseille, and some computer-generated.[2] The piece has three movements. The concrete sounds include ocean sounds, birds, insects,wood and metal chimes. Risset uses the computer in this piece mainly to process sounds rather than to sequence them. The sounds occur initially as "soundscape photographs", and then return as transformations, varying in proximity to their source sound. Risset uses several operations: filtering, modulating, reverberating, spatializing, mixing and hybridizing. He uses cross-synthesis to impart the dynamic character of one sound to another, for instance in giving the flux of the sea to different sounds. Another fundamental structural element in this piece is the use of major-minor pitch scale to color the natural sound timbres. This occurs gradually, strengthening in the last movement of the piece when sounds become quite dramatically fused into harmonic elements

---

[1]Risset, "Computer Music Experiments 1964- ", in Computer Music Journal

[2]from Risset,, translation of notes accompanying recording of Sud. For a more detailed discussion see "*Sud*, une musique numerique hybride et 'naturaliste', Bulletin du CPRIM

through the use of the scale as a kind of harmonic grid. He tunes the sea sounds, for instance, to a G sharp. Risset also uses spatialization to dramatic effect, placing natural and transformed images to suggest broad, open landcapes.

These techniques result in a work of great technical execution which remains firmly anchored in a musical experience. The subtle transformation of sounds and use of spatialization produce a feeling evocative of the beach, of human activity stilled. The piece conjures the mental state of a heat daze, or doze, where the distribution of sounds throughout space contribute to light-headedness.

Risset combines in his approach scientific, descriptive knowledge of sonic structures, the psychology of psychoacoustics and a background of musical composition. Like Harvey, he uses these facets to illuminate each other. The music exists as an aesthetic and a means to explore fundamental properties of human perception and understanding. He uses the computer as a tool to extract and articulate attributes of sound structure, and to "extend compositional processes on the level of this structure." The methodology of these composers, represents an exciting direction for an exploration of sound in relation to moving pictures which might also have such a basis in such a marriage of knowledge and expression.

How do sounds carry meaning?

The questions of how to imbed meaning in sound, and how to technically manipulate concrete sounds into musical form directly concerned creative issues in "Tales of Love and Glory". These include the articulation of content and my desire to avoid use of a spoken text as voice-over. I wished the work to suggest many layers of ideas, and I felt that the images contained much of the richness that could lead to interpretations of various kinds. The technique of adding voice-over, however delicately carried out, seemed too heavy handed to me, for even if done allusively, the voice directs the reading of the images. Here, I felt close to the cinema verite tradition: let the pictures and the sounds speak for themselves. Yet, I didn't want the piece to lose shape and impact because of vagueness about the intent behind the images.

I became more convinced that it might be possible to use the soundtrack as another film, the one the viewer wouldn't see. It would be an entirely sonic world, with ambient spaces, and occasionally familiar sounds popping out of less recognizable layers. Transformed sounds would almost come into focus as concrete familiar ones, particularly voice, but this voice would not emerge to explain the images. This voice would remain on the edge of intelligibility, the voice in dreams that is always heard but never speaks.

It became clear to me that the sounds would have to carry their own sense of meaning as a structural characteristic--as part of their inherent design. The events of this world would occur as transformations, the shifting of meaning and identity through the shifting of structural attributes.

My interest in the Wishart's ideas emerged because his conception of sonic images seemed 'cinematic' through his emphasis on context and inherent structural attributes. Wishart approaches the issue of scientific descriptive language for sound manipulation slightly differently than Risset and Harvey--he attempts to create structural categories as a means of ordering and understanding sounds. His method forms a useful, overarching set of considerations about the nature of sonic material.

Sounds in the world carry information about their origin through their structure. Sounds originating in a live recording carry information about their acoustic space, but what about the sound-objects themselves, as unique units? Wishart uses what he terms the "morphology" of sounds as one approach to structuring ideas about our comprehension of sonic materials. Morphology is a biological term relating to form and structure in organisms, but this term usefully illuminates a method of considering such inherent structural attributes as envelope characteristics or other features organic to a sound as a means of understanding and ultimately manipulating them.

Wishart carries through with some of the ideas presented by the Groupe de Recherches Musicales, who categorized sound objects according to how they continued in time. This resulted in three basic categories: the Discrete, the Iterative, and the Continuous. These types of continuation imply something about the amount of energy directed toward the material producing the sound, and also convey information about the nature of the material itself. Wishart observes that any "sound-event has an intrinsic and an imposed morphology." Wishart makes further distinctions in order to clarify how the imposed morphology, or the application of energy to the sound producing object of system, can reveal information about that energy and thus ultimately about gesture.

Gesture seems to a be feature of sound we consider in a performance context but less as an attribute of concrete sounds. Gesture in the latter has a musicality as well. The features of attack, sustain, and decay constitute natural gestural elements that are expressive and thus musically useful. Perhaps we understand this as the 'poetic' nature of sounds in the world. The attribute of loudness, especially in combination with an abrupt attack profile suggests force, power, violence. Continuously streaming sounds, such as bowed sounds,which emanate from a constant force of excitation, reveal something about the source of the excitation--that it is characterized by a kind of constant pressure.

Over time we can perceive changes in the amount and type of force applied to sound-objects which reflect physical activity. Wishart believes that qualities of imposed, gestural morphology and intrinsic morphologies function in two dimensions on our perception. Intrinsic morphology carries assumptions about the physicality of the source--not just source recognition, but perhaps an act of the imagination. Gestural morphology we react to directly because of our associations with human physiological process and the intellectual processing that occurs in our considerations of human activity. Thus, sounds with the latter features become anthropomorphised in some manner.

Structuring our examination in this manner begins to reveal certain relationships amongst sounds, which can lead us to another way of approaching a descriptive language about sonic properties. Maintaining sensitivity to the morphology of sounds becomes crucial to the success of synthesis or enlightened use of transformations. Computers allowed this kind of examination to extend far beyond the characteristics determined by the ear alone.

In the early days of musique concrete, before it was possible to easily analyze sounds, distinctions made by the ear governed compositional approach. Sound designers in the cinema still primarily work in this area of intuitive expertise. There now exist several different levels of research and methodology in the sound domain which attempt to unite scientific descriptive knowledge with an understanding of higher level cognitive processes. This suggests new areas of exploration for artists, such as myself, who wish to construct sound and image relationships in order to explore human imagination and the activities of the mind.

In TLG I wished to expand the associative space of the images through expansion of the acoustic space. I used various ambient environments to suggest that some other activity had a relation to the image other than the one illustrated within the frame. The power of sound to evoke space, through a host of tiny bits of data conveyed through reverberation and differential amplitude of sonic images offers a means to greatly expand upon the images contained within the frame. I have a sense of emotion about certain kinds of space that can be as powerful, or dramatic, as certain musical conventions, and I was curious in this piece whether this might translate. I began to think of the rooms of a house, each with its own particular ambience and perhaps activity--in one, a window is being opened, in another the drawers of a bureau pulled out and their contents examined. Could I also succeed in extracting bits of sounds and incorporating them into a musical structure, such that they would be perceived primarily as instruments?

MEANING AND INVENTIVENESS IN SOUNDS AND PICTURES

We have talked about various complex ways to order and manipulate sounds, and touched on the technology to bring this about. I have earlier suggested that the manipulation of sound in the context of motion pictures, both in the narrative feature film tradition and the more experimental endeavors in film and video remain bound to a set of technological conditions analogous to musique concrete. Having suggested a step to the more sophisticated tools and possibilities offered by the computer, we should consider the nature of the relationships between sound and image that might develop from this advance.

Pictures have content, which they describe or illustrate. Do they only convey meaning through the delivery of content? Earlier we have described a compositional approach which constructs aural landscapes and considers the morphology of sounds to carry various assumptions/kinds of information about those sounds. Sounds are surrounded by our ideas.

Can pictures which originate in the world--observed images--carry ideas beyond what they illustrate? This notion is important to my work since I chose in TLG not to use purely synthetic or abstract images, but images which contain photographic realism. Earlier we examined some ideas and techniques of manipulating sound in order to achieve certain psychological effects. There it became evident that the technology of the computer to digitally fix sounds and allow precise operations of transformations and ordering has vastly extended the power of the composer into areas of psychoacoustics and perceptual subtlety. How can we consider the effect of moving pictures in time in relation to these ideas about the apprehension of sound?

The interplay between the imagination we carry about sound and that which we bring to looking seems crucial to the exploration of a language of sound and image which does not rely on the spoken word or the conventions of narrative storytelling.

The importance of the photographic image

In terms of understanding the relation between technology and the obsession with time, we have to start, in the world of images, with a discussion of photography. The invention of photography, based in optics and light-sensitive chemicals, paved the way for motion pictures. The excitement about photography when it first appeared came out of its ability to trap time--it allowed access to instants. As cameras developed, the combination of the mechanical and chemical gave photography an authority which painting could not challenge. Cameras were "clocks for seeing[1]".

Computers give us access to instants as well, by digitizing them and by synthesizing them. The grain of photographic emulsion is made up of silver halide molecules which are light sensitive. The grains of the digital world are made up of zeros and ones. Since the images for TLG were primarily collected with a camera and emulsion-based technology, the properties of the photographic image are important to discuss both in their relation to my thesis work and as a frame for our thinking about digital tools. I suggest that the attraction of new technology resides not only in what it can reveal to us about the external world, but what it can teach us about ourselves. The history of photography is an example.

The fascination of photography resided in its power to represent the world, through the capture of time. Early pioneers of the photographic image, such as Nicephore Niepce, maintained as their sole object "to copy nature with the greatest fidelity.[2]" The images made by Daguerre as the first daguerrotypes were described as being "no longer transient reflections of objects, but their fixed and everlasting impress[3]".

> *....everything was reproduced with incredible exactness. The astonishment was, however, greatly increased when, on applying the microscope, an immense quantity of details, of such extreme fineness that the best sight could not seize them with the naked eye, were discovered...the effects of light and shade...were rendered with wonderful truth.*

In the entrapment of the moment we can perhaps discover or unveil secrets which we might not have noticed in the course of our daily activities, or if these moments had not been fixed for our scrutiny. This scrutiny is heightened by a kind of parsing performed by the camera. The camera does

---

1Barthes, Camera Lucida, p.15
2Newhall, The History of Photography p,20
3Ibid. p.23

not reproduce reality in full, it sections the world in front of it. The acts of framing and choosing camera position--decisions made by the photographer, result in a primary selection of subject matter which narrows our attention to that which lies within the edges of the photograph. The gaze in this watching has perhaps more direction to it than the gaze we employ in our daily activities since the latter has all the visual data of the external world to account for. While the photographic gaze is fundamentally one of realism, it is also one of limited focus on the object, the photograph, itself. But this channeling of our attention invites speculation about the world which lies within the frame. It is a vision that can occur over time in opposition to our scattered apprehension of the world as we process its varied manifestations.

The inventiveness of sight

> *(Muybridge's photographs) showed how inventive the eye is, or rather how much the sight elaborates on the data it gives us as the positive and impersonal result of observation. Between the state of vision as mere 'patches of color 'and as 'things 'or 'objects', a whole series of mysterious operations takes place, reducing to order the best it can the incoherence of raw perceptions, resolving contradictions, bringing to bear judgements formed since early infancy, imposing continuity, and the systems of change which we group under the labels of 'space, time, matter and movement'.* - Paul Valery[1]

Photography fixed fleeting moments mechanically, seemingly offering an objective, scientific representation of reality unclouded by the processing of human minds. Valery's remark above addressed the insight on human perception offered by Muybridge's success at fixing motion in time. If these images had not been captured, we would still not know the truth about how living beings moved through space--we would only know how we imagined them to move. Photography teaches us about ourselves as it captures the world. Artists, those individuals with the responsibility of presenting such images to the world, operated under what Valery calls the "law of unconscious falsification". The images they presented to the world created

---

[1]quoted by Frampton in Circles of Confusion ,p.78

conventions in their times about appearances and behaviors. Their renderings of the world were not untrue, they were just inaccurate, processed by the embroideries of the mind.

What happened to the inventiveness of sight? I don't think it disappeared with the increased realism of photography. Valery is correct in that photography showed us something about ourselves; the inventiveness of the human mind. But photography did not stop the mind from elaborating on the data of the image. The imagination occurs as a different order of phenomenon, on merely a different level of inventiveness. We can regard a photographic image and still have many imaginings about what that image contains for us. The 'law of unconscious falsification' still holds sway. The scientific objectivity of photographs so applauded in its early days does nor prevent viewers from projecting a bit of themselves into the frame, from embroidering the image with all sorts of ideas about the subject, the photographer, what occurred just outside the frame, what the conditions were surrounding the taking of the image. This is a powerful phenomenon to remember, not just in the context of photographically-derived images, but in the general context of image-making.
From the scientific objectivity of the photograph we have come to an understanding that all these images have a point of view, the frame of the photographer. Beyond even the eye of the photographer is the eye of the viewer. However impartial this manner of collecting images may seem, the frame gathers and focuses attention on a collection of aspects of the observed world in a way which encourages the viewer to have ideas about that world.

The relation between recognizable visual elements and the imagination became quite clear to me as I consider the use of a collection of processed video materials in TLG. During the course of making this piece I became intrigued by the possibilities of real-time video processing offered by the Fairlight Computer Video Instrument. This device operates by having a small quantity of image memory, called a stencil plane, and various features which allow manipulation of the image. Video can be grabbed as a whole frame, drawn out to reveal unprocessed video underneath, combined with artwork created in the stencil by paintbox techniques, or built up in the plane through operations such as strobing, where the video is grabbed according to luminance amplitude at a certain rate and stored in the stencil plane sequentially.

Despite the Fairlight's sometimes clumsy features, I liked the look of the process, which can be quite thick and painterly, and that the process originates in a real image. Thus there is a kind of parent-child relationship between the images, which I thought had particular relevance to this project. I also liked the allusion to memory in this process--that images are built up and encrusted in the stencil plane somewhat the way experiences become layered in our memory. They partake of the original experience, but our imagination is also at work, transforming them into entities with their own shape and texture. I processed quite a lot of the footage, particularly the first, documentary materials.

Originally, I thought that the thesis work would involve quite extensive use of this processed material, and would thus investigate the use of digital tools and both sound and image. In film, any kind of image processing must proceed through painstakingly slow optical procedures, so I was quite intrigued by video's accessible plasticity.

Ultimately I have not used very much of the processed imagery in "Tales of Love and Glory". I found that over time the visual quality of the material became very flat, and did not necessarily integrate well with unprocessed images. As an editorial choice, the manipulated images seemed increasingly formal--the processed images did not retain any emotional quality, which is highly problematic. They become distancing, a retreat into process instead of a revelation through one. The images lacked the suggestiveness of those images based in the real world, mainly because it was so hard to control the balance of transformation. Compelling possibilities exist for the use of image processing tools that can handle large quantities of image data, but they must also have the processing capabilities to mirror the kinds of subtle processes we have discussed in computer music research.

THE PROCESS OF CREATING A UNITY OF SOUND AND IMAGE

Montage and synchronization

Two essential concepts to consider in this discussion are montage and synchronization. Montage, the arranging of picture elements, or shots, in time, is an aspect of motion pictures which has led to extensive theorizing. It has importance here because of its parallel in musical form through rhythm and duration, and because "Tales of Love and Glory" is structured primarily by the straight cut. I have not used techniques, such as extensive image layering, keying or blue-screen which are considered to belong to video, because they can be executed electronically rather than optically, and are therefore arguably more accessible. This piece is based primarily in the language of cinema in its dependence on the cut. I have used this to create links between these otherwise disparate materials--images shot at different times, in very different locations. The relationship of these images to each other can be articulated by the quality of similarities or contrasts to each other, by aspects only viewable by contrast with the image that follows.

Ideas about montage tended to incite polemic theorizing in the early years of cinema. We are now so bombarded with images presented in so many different modes of delivery that the distinctions fought for by these pioneers seem extreme. But these distinctions remain useful to consider as a set of tools, and a way of looking at what the effects can be. Eisenstein believed that shots related to each other dialectically, through a set of 'collisions' which produce a new idea.[1] He established this idea in reaction to other current beliefs in the Soviet cinema, that of Kuleshov's "relational editing", and Pudovkin's belief that montage controlled the 'psychological guidance' of the viewer[2] . Eisenstein made a distinction between linkage and collision, which nows seems a bit extreme. The great potential of cutting arise from the presence of both of these effects through the joining of images.

---

1 Eisenstein produced a large body of theoretical material on the cinema, but two of his important essays on this topic are: "The Cinematographic Principle and the Ideogram", and "A Dialectic Approach to Film Form", both are found in *Film Form*, Essays in Film Theory, J. Leyda, ed. Harcourt, Brace, Jovanovich, 1949.

2Monaco, How to Read a Film, p. 309

Eisenstein's notion of the shot as ideogram relates quite strongly to the notion of the sound object. Eisenstein conceived of shots as self-contained units of meaning similar to the use of pictorially-based language forms in oriental languages--specifically Japanese. The meaning of the shot was a structural attribute, inherent to it without the addition of any external factor. Meaning was built through the collision of these shots, or montage cells. But just as the sound object has attributes which vary over time; the envelopes of amplitude or frequency, so individual shots have properties which vary over time as luminance, color, and motion. This method shifts emphasis away from considering montage, or editing, as narrative explication and stresses the relation between editing and an activity of the mind.

We talked about the activities of Murch, Splet and Kubelka in the context of musique concrete. Kubelka unites this editorial approach typified by Eisenstein within the domain of sound and image relationships in the context of experimental cinema. Kubelka has produced some extremely rigorous works. However powerful and creative Splet and Murch are as sound designers they remain tied to someone else's vision. Kubelka represents another kind of filmmaker who will not attain the fame and huge scale of these two but who can remain in control of the entire process of composing the sounds and pictures which make up his films. His output has not been large by any means, but in addition to *Unsere Afrikareise* he has produced the films *Adebar, Schwechater* and *Arnulf Rainer*. Kubelka manages to achieve a high degree of inventiveness although his materials remain quite firmly rooted in the concrete. In *Unsere,* he does not use synthetic or heavily processed sounds, but creates an aural landscape and a series of very precise links through synchronization which begin to become almost unbearably suggestive.

Synchronization as an extension of cinematic language

Synchronization makes links between the sound world and visible events in the picture. A film such as *Unsere Afrikareise (Our Trip to Africa)* brings out the extent to which synchronization can be used a form of address. This film started as a commission for Kubelka--he was asked by some Germans heading to safari in Africa if he would shoot a film of their trip. He recorded about ten hours of sound and a few hours of image[1]. The sounds included animal noises (this had been a hunting safari), talking, radio, and ambience. Upon his return he transcribed all the sounds phonetically, and developed a technique for scoring passages for some of these sounds. He clipped three frames from every shot and mounted these on cards. Then, he cross-referenced these according to rhythm, color, subject, and theme. The totality of these images and sounds he designated his 'vocabulary'.[2]

Kubelka believes in re-conceiving a film completely after shooting, and he spent six years re-making his materials into the finished, 12-minute film. Crucial to his final language here is what he terms, "synch events". These are highly determined conjunctions of sound and picture events, none of which were present in the original context of shooting. By using such precise indexing and analysis of his materials, the basis for the re-constructed synchronization could work around many sorts of elements in the picture: color, content, motion in the frame. Thus, a passenger boat riding a crest of a wave is accompanied by a rising swell in a snippet of music, an overlapping which produces an emotional surge. A shot of insects associated with a more lively musical scrap suggests an insect dance, later, this same music appears under a shot showing black Africans struggling to lift the carcass of a dead lion onto the top of a vehicle. Here, the the rhythmic heaving of the men as they grapple with the heavy carcass is in as perfect synchronization as the dancing insects, but the effect is forbidding.

By carefully shifting around these "synch events" with a limited picture vocabulary, Kubelka arrives at a discourse about colonialism, tourism, race relations, violence and the offhand treatment of the natural world by modern man. There is no voice-over, and in fact, what little recognizable language exists imbedded in the tracks contains no English. I have never "understood" the film in that sense, as most of these language fragments are in German, a language I do not know. But the discourse is both very

---

[1] Sitney, Visionary Film, 1979, pg. 302
[2] Ibid.

complex, indirect by definition because it has its basis in post-synchronized sound, yet completely explicit.

The technology used by Kubelka is the traditional magnetic based media of the 16mm filmmaker, and the aesthetic that of musique concrete and the collage. The subtlety of the execution resides in the extreme care taken around the placement and juxtapositions of these materials, and the mixing, which brings the elements in and out of clarity, or focus, on the track.

Here, the power of synchronized sound events to picture lies in the mental associations which get built, then shifted, and built again with a degree of such surety against the image that the result seems utterly convincing as "reality." In its conventional use, synchronized sound contributes to the sense of authenticity--the "reality"-- of the film by its perceptual affirmation of causality. The actions we see are accompanied by the sounds we expect to hear from our experience of the world. The affirmation of this expectation leads to a certain belief in the image as a credible representation of the world.

The reversal of the expectations created by sync can become a source of ambiguity and surprise in a manner reminiscent of Harvey's sonic transformations. Our perception of these events becomes freshened, re-oriented according to a new set of data which makes connections between unexpected objects. This results in a very active process of viewing.

In "Tales of Love and Glory" a sense of ambiguity was necessary in order to accommodate the various layers of interpretation that existed within the images. The scenes that take place in the New York City playground suggest violence, anger, isolation and freedom. They evoke a certain grace of childhood, yet also its fragility. The feeling of the playground there is somewhat sinister, a suspicion which becomes confirmed by a battle between male and female teenagers, where the boys humiliate the girls by breaking raw eggs on their heads. The emotional edge of all the images of childhood remained constantly present to me, and it was important that the associations of these images carried over to the images which were more concerned with the adult world. I used synchronization as way to make some of these bridges, as a way of suggesting certain relations. However, all of these synchronized events were fabricated. I also found that synchronization with picture elements created a more immediate experience of the image, and was important in constructing an overall dynamic shape in time. The precise relations synchronization constructs bring viewers into a kind of attention to the image.

Each shot in "Tales" represents a thought, an idea. The discourse becomes one related to dreaming, or poetry. It is a discourse based on what the images depict but also, as in our examples with Harvey's treatment of the boy and bell sounds, the constituent parts of the image. For the process of finding sounds for these images requires a deep analysis of visual qualities which have some kind of sonic mirror.

Sound and Image as formal relations

An examination of the formal elements of picture motion over time brings us back to music, this time less in the aspect of montage, the changing of shots over time, but in the consideration of changes within form over time. We might understand this an attempt to articulate formal attributes of the picture. The origins of this direction of formal exploration in the cinema trace back to early color and music experiments, notably realized in this century by the composer Scriabin. He had a 'color organ' built by Alexander Wallace Rimington, which projected colors onto a screen to accompany his composition of *Prometeo,* in 1911. The scheme here was fairly simple, where each color corresponded to a note on the Western scale[1].

A handful of filmmakers in the early days of cinema became interested in extending musical ideas of color, rhythm and form into the time-based expression of cinema, specifically Viking Eggeling with *Symphonie Diagonale* (1921) and Hans Richter, who produced a series of abstract

---

[1]Abbado, "Perceptual Correspondances of Abstract Animation and Synthetic Sound", MS thesis

animations between 1921 and 1925 called *Rhythm 21, 23, and 25.*[1] These films were done as silent films, yet curiosity regarding the musicality of the cinematic image fueled these works. Oskar Fischinger continued these experiments, adding sound. Len Lye developed techniques for image-making directly on the surface of the film in the 20's, and by 1935 had created *Colour Box,* with music. Norman MacLaren developed a very precise methodology for working with abstract shapes and sound.[2] These explorations into musical forms and cinematic ideas explored such relations as that between 'tone colour' and visual color or forms, or the rhythm of color changes on film as an analogue to rhythm in musical composition. Many of these techniques involved directly affecting the film through painting techniques and abandoning the camera altogether. For a detailed historical overview of this activity see [Abbado88]

In this scheme, for instance, timbres with a strong distribution of high harmonics would fit well with images with sharp edges, high reflectance, or 'bright' color. The formal exploration of this particular area does not deeply interest me at this time. But there are aspects in this approach which can be quite instructive. We can observe tangible relations between the physical attributes of objects and distinct musical elements. The awareness of relations between tone color and visual color offers opportunities to experiment with timbre selection and synchronization manipulations. The matching of surface attributes of the image to the harmonic structure of timbre can render a compelling effect of unity, which is a powerful effect. The synchronization of these effects becomes powerful in the manner described above with concrete sounds. The establishment of a precise relation even of abstract properties leads to a conviction that *these* sounds belong to *that* image, even if picture and sound are completely invented.

---

[1]see Rusett and Starr, Experimental Animation, Van Nostrand Reinhold and Co. 1976 for detailed discussion of the work of Eggeling, Richter, Ruttman. Fishinger, Lye and MacLaren.
[2]see "Norman Maclaren and the National Film Board of Canada", in Russett and Starr, eds. Experimental Animation. pg. 116

In "Tales of Love and Glory" I have had to consider some aspects of the image on a fairly formal level, despite the fact that the images are not abstract, but generally concrete. This fine level of consideration--of treating the formal, buried features of the image as well as the more obvious features related to content, becomes an important source of ideas about the nature of the sounds which might accompany the image. This process of dissection enables the construction of a sort of palette where all the potential elements can be examined. The colors, luminance values, grain, or noise within the image, the patterns of movement within the frame, or the static, graphic compositional qualities of the image all become elements which can have a sonic dimension. The motion of the camera, the use of screen direction, either as the motion of objects or the direction of actions or glances within the frame also can become part of the dynamic. Here, we can see how montage becomes important as a means to articulate these aspects of a shot.

"Tales of Love and Glory" incorporates both black and white and color imagery. I use the color for a certain kind of emotional exploration relating to light color and dramatic perceptual moments. Most of the shooting of the "childhood" images occurred in the fall, when the light is very distinctive and evocative. The color material also expresses a sense of direct emotional experience, it has a quality of the "present". The dramatic surprise of the vivid green color of the grass, and the tremendously subtle colorations of the water as it explodes into abstract forms (see fig. 18) are all related to kind of privileged apprehension of the world which exists in contrast to the images which observe the world in starker terms. Colour within the image, and the texture produced by grain/noise, contrast offer a means to structure certain ideas about timbre.

The set of ideas which concerns purely formal relations represents an aspect of the possibilities contained within the idea of sound and picture correspondences. I see it as part of a foundation. The work already executed by Fischinger, MacLaren and John Whitney stands as excellent explorations of the relationships between properties of abstract objects and music.

Today we have the technical capability to push both the imagery and sound elements farther--we can use computers to more accurately model shapes, and sounds, and to control their properties quite precisely. But it seems to me that we should use this sophistication to push our explorations into areas of deeper psychological complexity, not just formal complexity.

Towards a cinema of the mind: poetics

By that, I mean work that probes into that part of us which processes meaning, emotion and memory. These areas of our psyche allow us to have feelings and ideas about what we see and hear in the context of film, and in that way to form a connection to the work which can be, at its best, powerful and illuminating. We are at a stage in the development of tools for our expression where the level of affect can become much finer due to enhanced access to the discrete components of the material. It may thus become possible to make works which can address our perceptual system very directly.

One hopes that the viewer will have a "meaningful experience"--how to involve him with the material so laboriously collected and shaped? One way might be to make the perceptual experience itself involving in as profound a way as perception itself, exploiting our mind's ability to make meanings and derive connections, to touch more directly into the perceptual basis of the experience. Risset expresses this hope for the future of computer music, and I would add that it presents the challenging direction in developing sound and image relationships:

> *Evoking a suggestive yet illusory world, free of material constraints, by playing directly...upon perceptual mechanisms, thus unveiling perceptual primitives and guiding perception toward one mode or another (e.g. synthetic vs. analytic)*[1].

Risset works entirely in the medium of sound, where images are evoked but not illustrated. The addition of picture to sound which is highly suggestive and imagistic in its own right has to occur with the proper balance or the

---

[1]Risset, "Computer Music Experiments, 1964- ", Computer Music Journal

two streams of expression may appear too disconnected. This raises questions about a mode of delivery which can lead to a true integration of sonic and visual materials. The form which will allow these two powerful modes of expression to complement each other is not a familiar form of narrative convention or linear construction, but rather one which echoes the human processes of deep mental activity and dreaming. In the mode of dreaming, particularly, we find that the fragmentation of time, the layering of sensations and ideas, all achieve a substance and comprehension which the mind would otherwise dismiss as confusing in states of waking consciousness. The stimulation of dream-like experiences in the wakeful state of consciousness becomes additionally powerful due to the conscious processing of such experiences.

This question of form then begins to specifically address questions of cognition rather than narrative development in the classic sense. The deviation from such a deeply rooted convention places the onus on the artist to articulate the new form. Andrei Tarkovsky, director of such films as *The Mirror, Nostalghia,* and *The Sacrifice,* uses the term, "poetic" to describe a cinema which obeys its own laws to touch human , and not those of literary or theatrical traditions. He suggests that "poetic reasoning is closer to the laws by which thought develops, and thus to life itself, than is the logic of traditional drama." To him, poetry is not a genre, but "an awareness of the world, a particular way of relating to reality." The poet is "capable of going beyond the limitations of coherent logic, and conveying the deep complexity and truth of the impalpable connections and hidden phenomena of life."

Tarkovsky struggled, in his work, to see "life below the surface." The value of artistic activity lies in this search, and my own work concerns this above any other technical or formal value. The tools exist merely to make this search a more compelling one to share, and a deeper one to explore. My conviction about the expressive power of the computer in the sonic dimension, and perhaps ultimately in the visual one as well, is the extent to which it allows us to probe beneath the surface, and seek out these connections. This can not occur as a purely mechanistic operation. To return to our examples in the sound dimension, of *Mortuos Plango* and *Sud,* here it is evident that the sensitivity of the artists to their materials, their understanding of their tools and their process, and their curiosity about the nature of "connected-ness" bestows their work with this sense of the poetic which Tarkovsky describes.

> *We have these two ideas: the belief that dreams are part of waking, and the other, the splendid one, the belief of the poets, that all of waking is a dream.*[1]

Borges, in the quote above, suggests that the poetic arises from that aspect of ourselves which doubts that reality is anything other than the mind itself. If we consider poetic then as the creation of a mental activity, the expressive possibilities for the poetic use of moving pictures and sound become more clearly oriented to the stimulation of a state of mind and less obscured by trivial interpretations of the term itself. The "Poetic" use of cinema often teeters dangerously toward the self-conscious and artificial. To achieve a poetic use of sound and picture, one which operates on the level of thought, the mind must find itself fully engaged in the act of mentally processing the arrangements of sound and picture objects over time.

Film and music, because they incorporate elements which occur over time into flexible structures, are the most appropriate media available through which to explore those aspects of our minds which are not restricted to linear, analytical reasoning. The tools of montage, synchronization and an informed sonic vocabulary allow us to play with our apprehension of the world, instead of carefully and faithfully re-creating it.

Because of the parallels between a cinematic experience and our waking consciousness, the medium of motion pictures immediately proved itself to be a powerful tool for exploring the unconscious. The process of making movies is fundamentally one of recording an event (either observed by, or created for, the camera) in the real world and and recreating a representation of that event in the mind of a viewer. The viewer knows that what the is seeing is not 'real', yet the experience can have a similar

[1]Borges, J. from "Nightmares", in Seven Nights , p.30

immediacy. The life of the mind overcomes the life of the body. The earliest screenings of films made that effect very clear, as in the Lumiere's *L'arrive d'un train a la gare du Ciotat.* This film is very brief, one of the first ever made, and shows simply a train arriving at a station, shot from the platform in the direction of the oncoming train. Early screenings apparently produced pandemonium in the movie theater, as the train pulled into the station seemingly on top of the audience.

The only other times human beings have this experience of the life of the mind completely overcoming that of our corporeal selves is when we are dreaming. The surrealists caught on to the connection between cinematic language and the language of dreams and provocative psychological experiences and exploited it in films such as *Un Chien Andalou* (Bunuel, Dali, 1928), *Anemic Cinema* (Marcel Duchamp, 1927), and *Etoile de mer* (Man Ray, 1928).

These early explorations into dream-like, associative filmmaking were quite prescient about how naturally and powerfully the medium of film might explore human perception. Though the films mentioned above were silent, their use of fragmented time, conjunctions of objects within the frame and between shots, surprising and sometimes violent elements in combination with the banal and familiar, to evoke an experience of psychological complexity, hold up to any contemporary work.

This work remains powerful because these artists had discovered a way to make the work "meaningful"--it perhaps did not make traditional sense but a viewer would not walk away untouched, either. This is one way for such work to have meaning--if it has some distinction from the plethora of sights that occupy our lives. An image does not become meaningful through purely formal elements. One experiences it as meaningful if it contains elements that one can process into meaning. To view a work such as *Un Chien Andalou,* or *The Mirror,* is to create a very personal set of ideas about what the film means and what it is about. As the viewer devises his interpretation of the work, he integrates his own life experience, thoughts and beliefs, and through this integration the work may come to represent something significant to the viewer, as an event which catalyzed this synthesis.

"Tales of Love and Glory"

The above considerations became important to the process of making "Tales". I wanted this work to have "meaning" for viewers, that elusive word, but of the subtle, imbedded kind that we earlier described as poetic. It seemed to me that a successful approach might be one where sounds and images served as material for a certain kind of mental process. The mind seems inclined to make connections, order objects, search for patterns amidst chaos, derive meaning from perhaps random elements. The world is full of sights and sounds that may or may not stimulate such mental activities, but if these sights and sounds are re-presented to us in certain ways they can become quite stimulating. Artists know that through re-contextualizing the familiar a host of new ideas can be generated. The mind makes inventions out of everything it sees and hears, and by running picture and sounds through time more kinds of invention are possible. This can be a re-contextualizing of objects within a frame, or one that occurs between "frames", as in montage. Montage, used powerfully, can capitalize on this to defy expectations and activate associative mechanisms. I began to consider music as being an extension of montage, in the way it arranges bits of meaning in time, either in the form of patterns or concrete sound images.

I found that the tension between visual objects that were mostly recognizable and music that was sonically dense and unfamiliar contributed greatly toward the kind of associative space necessary to drive TLG. If we think of the work of Escher and Magritte, the unsettling quality of their images comes out of the juxtaposition of the familiar with the incorrect. All the usual visual cues are present--colors, forms, edges, surface renderings, shadows, but the overall relationships of the elements do not add up to the reality we expect. That feeling of being on the edge of something totally familiar yet completely unknown remains profoundly intriguing, because of the proximity we attain toward our own perceptual workings. We tend to call this ambiguity 'dream-like' or hallucinatory'.

Unlike Escher and Magritte's images, the visuals in TLG unfold in time. Through duration transformations inherent in the image may emerge, or certain points which become highly suggestive. Barthes has a useful term here, 'punctum[1]', which he applies to still photographs but I think applies to the moving image as well. In Barthes' applications of the term the punctum is that detail which not only attracts attention, but "its mere presence changes my reading"--it is a "pricking" which changes the "value" of the photograph. The punctum is that visual element which elevates the image from the commonplace to the status of 'meaningful' as we earlier discussed. By maintaining representation to the world, the familiar, the expected cues are laid out. Then, the experimentation occurs with how to turn around expectations through the joining of shots. Could we imagine an Escher image that built up over time, whose power didn't reside in one frame, but over many?

> *We propose another, radically different morphology...one that views film, not from the outside, as a product to be consumed, but from the inside, as a dynamically evolving organic code directly responsive and responsible, like every other code, to the supreme moderator: consciousness.*[2]

The developments in cognition and the tools to access structural attributes of sound and image suggest that we can come yet closer to that "supreme moderator", which holds so many mysteries for science, art, and human understanding.

---

[1]Barthes, Camera Lucida, p.27

[2]Frampton, Hollis. Circles of Confusion

SEARCHING FOR THE RIGHT TOOL

The search for a more powerful set of tools in the sound domain led me to investigate the use of a computer-based system. The aesthetic motives driving this search have been previously discussed . The present section addresses questions of interface and design in existing systems, particularly the New England Digital Synclavier system and the WaveFrame Corporation's AudioFrame, and poses some issues for general consideration in terms of integrated audio workstation design for film and video post-production.

> ....*When paradigms change, the world itself changes with them* [1].

Is it worthwhile for a video- or film-maker to 'change paradigms'--to abandon the old and familiar tools for the new world offered by computers and digital technology? Obviously, it's worthwhile if the tools help to realize ideas. Kuhn points out in the context of the above quote that, "Led by a new paradigm, scientists adopt new instruments and look in new places." This applies to our situation as well. There is a sense that much more might change than just familiar ways of executing tasks. Technology gives us opportunities and obliges us to explore. In this part of the discussion we will examine interface issues brought up by these two systems. We will favor the viewpoint of film and video post-production over composing, per se, but not omit interface problems relating to the activity of making music.

Constraints

The handling of sound within the film/video environment has built-in constraints that shape the needs of the user in specific ways. Time manifests itself in a practical manner, in the form of synchronization.The needs of this hypothetical user, whom I will call, purely for clarity in this discussion, the sound artist, do not necessarily differ from the needs of a composer. He, too, needs to place events in time, but the details --of time-code standards, laying back to picture, originating timing data from picture cues, editing within strict timing parameters, defining sections of audio as shots, integrating edit decision lists with sequencers and external picture edit decision lists--form

---

[1]Thomas Kuhn, The Structure of Scientific Revolutions, pg. 111

quite a distinct set of specialized needs. I believe that any system that sets out to be useful within the context of film and video post-production must at a minimum be suitable for these needs. The capabilities of sound modeling, including filtering, analysis and synthesis, are, or should be, as important to the sound artist as they are to a composer. The workstation environment should not place either user in a ghetto, but hopefully offer an integrated environment where needs can be satisfied without too much compromise of ideas.

For a film- or video-maker, the activity of composing music or indeed, any kind of sound manipulation, is constantly dominated by the presence of the picture. There are certain specifications about the picture that will undergo many changes until the picture is 'locked'. Shots change length, move around or disappear entirely. A system must be able to handle the unstable and stable periods of project development. Decisions, judgements and gestures in the sound have to be evaluated with the pictures they go with. Often the unexpected occurs --sounds that one thought would be 'perfect', don't work at all for a host of reasons. The tracks, on their own, might sound wonderful, but when put against the picture they conjure the wrong mood, the wrong associations. Perhaps the nature of a particular sound and picture juxtaposition introduces a host of questions about the intent of the maker that don't belong to the act of viewing the work. Any of these situations can result in the artists' abandoning a creative masterpiece. The process is both highly open-ended and ultimately constrained. The quality of the relationship of sound to picture remains the final judgement on the usefulness of the sound, not the production value of the sound itself.

Because of the constant picture reference and the changes it often goes through in the course of designing sound, sequencing through the computer can offer some terrific advantages to the sound artist. A computer allows the artist to make sequence drafts following a word processing model. This could include sound effects and music data. Several of the well-designed and fairly inexpensive versions of sequencing software available run on inexpensive machines. Performer™ and Q- Sheet™ are examples of software that runs on the Macintosh. Thus, even a fairly simple set-up can now offer a degree of flexibility and experimentation within the overall constraints of picture editing decisions that could not exist previously.

It is also easier for the artist to flip between musical and effects editing activities. Although this flexibility goes against the Hollywood model of compartmentalized tasks, it has advantages in determining when effects and music are stepping on each other. One may not want a cymbal crash to occur over a door slam for instance. Ideally, the working environment should allow the composer and the designer access to each other's materials, at least as a reference. It should also allow for more fine-grained solutions to problems such as the one above--that is, in the conventional scenario, if sound elements conflict with each other, they can be either pulled out, reduced in volume in the mix, or equalized to occupy a slightly different spectral area, which may offset the sound just enough from the conflicting elements to allow both audibility. A potential advantage to computer-based systems is the ability to achieve modifications on the level of spectral composition, through such tools as spectral analysis and re-synthesis.

As in any creative endeavor, the artist needs to feel free to play with her materials. At the very least, her tools should not restrict her to the point of severe compromise. In my case, I knew I would need to move elements around frequently, and I also knew that I wanted to be able to experiment extensively with all of the sonic materials, from the smallest unit to the overall structure. Practically, I would have found this project terribly cumbersome to do with a multi-track tape recorder. Conceptually I am sure that I would have experimented less.

After all the decisions about order, sound design and composition have been made another potentially onerous stage of a purely technical nature must be executed. All the elements must be married into the final release media. In film, that specifically means mixing all the working elements down to masters for optical soundtracks, or for six-track Dolby™ release. Both of these require different processes and considerations. In video, the tracks are generally mixed down to the audio tracks on the master 1" videotape. This stage represents the final series of operations and usually the culmination of much time and labor. It is a stage at which maintaining the quality of the audio materials becomes critical, because the decisions made at that point are final. A workstation system must be able to support these activities. This stage is characterized by long playback periods, as the entire program, or large chunks of it, are mixed. It may be a repetitive process if questions exist about relative volumes or filtering. Complete accuracy in synchronization and dependability become crucial here. Fine adjustments of level and equalization may also be useful. Traditionally this stage would occur as a formal mix-down process, but I think we will see a trend toward accomplishing this through an integrated workstation in certain areas of production that require fast turn-around, such as broadcast.

All of these stages represent areas where new tools could offer solutions to problems that are purely practical in nature. I feel that the area of audio production for motion pictures is poised practically and aesthetically for a "paradigm change". Earlier attempts to introduce advanced audio technology, such as the Soundroid project, suffered through timing in the technology base and readiness of the industry to accept change. The momentum of the technology now seems strong enough that potentially large-scale change is closer. It is therefore timely for future users to become more concerned with the direction of computer-intensive technology.

Tools are propagated in cultures not by their inventors but by hapless individuals, such as myself, who struggle to get something done. The tool gains acceptance if this something is a goal shared by many people, and if the tools works. A feedback relationship exists. The user must get some benefit from the tool, the tool addresses a need. In another time, if you didn't like the heft of a chisel, you could take it back to the blacksmith who made it and ask her to re-forge it. In time, if enough customers wanted similar changes, the blacksmith might adopt these modifications into her tool design. We are in an era of proliferating, technology based tools, where this loop is expanding as the number of products and consumers grows. Makers become distant from their consumers. Consumers become distant from their needs, bewildered by all the options before them which almost answer their needs, but not quite.

The changes in technology that face us today often impact more than just the user. Computer-based tools tend to replace people, not other tools. So, advances in this area can seem pretty threatening to the very individuals who are also potential users. I have a strong belief in the importance of bringing tool-makers and users together, as I foresee unfortunate consequences unless we all participate to clarify the shape of progress.

There are many individuals making the transition from their familiar tools to a new generation of computer-based tools. I suspected that a better tool existed for the job I wished to do, but it wasn't easy to actually find one. My searching led me to realize how few tools in this domain were designed by people who really knew the nature the job they were purporting to address. There is a confusing array of choices currently available in this area, and faced with that, a user has difficulty focusing on her specific needs.

The nature of physical objects helps clarify needs

The formats of film and video lead to somewhat different needs because they are so different in nature as physical objects.Articulating these properties will help clarify important design issues.

Film is concrete--sometimes a little too much so, as when you can't find those crucial three frames, or find yourself dropping the core out of a 1200' reel. But there it sits, right in front of you. In the case of sound, I find that I begin to recognize and conceptualize my soundtracks by their patterns of fill, track and leader. This objectification helps me plan local tasks and global decisions about track layout. The sound elements are both sounds and objects--bits of mag track in certain sprocketed lengths. The sprockets provide one with a tangible way of measuring and comparing--is this length of sound as long as that shot? The task of checkerboarding tracks--of arranging tracks with the correct amounts of sound and blank leader-- is not an abstraction in film, but has a gratifyingly real visual correlation.

The world of electronic images seems very slippery in relation to the 3-D comfort of film. Images live in some baffling way on videotape, which offers no clues to its contents without the appropriate machine to decode it. To make matters worse, the sound lives there, too, which means sound and picture get married right off. This is a peculiar notion to a film person, who lives with separate sound and picture tracks until the most final stages of project completion. Laying out tracks in film, while cumbersome, allows frame-accurate control over placement. You can write notes to yourself, or someone else, on the back of the mag track.

It is possible that I would not have become so interested in exploring a computer-based tool for audio if I hadn't come from a film background. That the relationship between filmmaker and material is so grounded in physical attributes contributed to my thinking that computers could restore a visual object-ness in an area where it felt conspicuously lacking. I was so accustomed to a direct, tactile experience with my materials that I found myself often frustrated by the distance imposed by video. A film editor has a very physical relationship with the film he is cutting. I started thinking about computer interfaces in this perspective. How could the interface restore the feeling of a physical object into what had become a process abstracted from the source materials? The electronic, magnetic environments of video and multi-track audio recording leave me in a cognitive void. How can I have a relationship with seemingly identical rolls of brown oxide, that I can't even write on?

The successful examples of software in this area understand and exploit the need for visual cues by devising visual representations which the user can move around and do things to. MacMix[1], a sound mixing program designed by Adrian Freed, offers a good example of a software emulation which allows the user to execute familiar tasks in software without having to re-learn them in the terms of the machine. It's a transparent switch. A professional mixer or sound editor who has spent years amassing certain skills, will not want to be forced by a system to radically re-learn or modify her technique if it means merely coming to grips with how the system sees the world. If the discrepancy is that wide, the fit that off, between the user and the new environment, this new world may not be designed by someone who really knows what's important in it. The sound editor or mixer is probably right in avoiding it. She will change paradigms if it does not mean throwing away too much previous knowledge and if the benefits really offer a considerable advantage over the old way of doing the job.

---

[1]Freed, A., and Goldstein, M., <u>MacMix 1.2A.</u>

> *Use technology to make visible what would otherwise be invisible, thus improving feedback and the ability to keep control*[1].

Developments that push the technology

The audio domain of the film and video industries seems poised for some big changes, due to certain recent innovations initiated by the infiltration of computers and abetted by other factors such as MIDI standardization and fast, inexpensive DSP chips, such as the Motorola 56000. Computerization came to the motion picture industry first in the visual domain, in the form of edit controllers and special effects devices. Video production has been quicker to incorporate these advances than the film-based branches of production. But audio has still been caught up in fairly conservative technology, based on the multi-track tape recorder and magnetic recording media. The film industry remains perhaps the most conservative of all, holding on to the magnetic/mechanical tradition and a large hierarchical labor force to make it suitably productive. Here the union structure also inhibits innovation. But it might be worthwhile at this point to consider innovation and revolution on a somewhat broader scale:

> *Why should a change of paradigm be called a revolution?...scientific revolutions are inaugurated by a growing sense, again often restricted to a narrow sub-division of the scientific community, that an existing paradigm has ceased to function adequately in the exploration of nature to which that paradigm had previously led the way. In both political and scientific development the sense of malfunction which can lead to crisis a prerequisite for revolution.*[2]

Will we see a revolution in sound production in the film and video industries? The existing rules have not changed, but perhaps the exploration of "nature" has, and the current paradigm can no longer lead the way. A glance at industry publications such as Mix magazine and Music and Technology reveals many articles about how film and video post-production facilities can adopt the computer-based technology which has evolved through

---

[1] Norman, The Psychology of Everyday Things

[2] Kuhn, T. The Structure of Scientific Revolutions.

music production. This set of production values is rapidly setting the standard for audio quality through its ubiquitous presence in the commercial music marketplace. Increasing numbers of film scores for the feature film industry are being produced through the technology of MIDI, sampling, and digital production techniques. The other aspects of sound production will eventually have to keep up this level of quality at the very least. Hopefully , we will see a greater leap than mere production value, one that will tie into the expressive potential of a computer-based sound environment, as earlier discussed.

History of the technology

> *I believe the key to the future of computer music is the digital audio processing station*[1].

The search for a unified digital environment for the manipulation of audio materials has fueled much research in computer music. Moorer points out that any general purpose computer both affordable and sophisticated enough to be useful to a composer would still be too slow by two orders of magnitude for real-time synthesis[2].

IRCAM had recognized the need for specialized hardware to accommodate the processing requirements of audio synthesis and processing in 1978. The 4B, 4C and 4X machines designed by Guiseppe diGuigno were directed toward this problem. The 4X, used for the synthesis and processing of natural sounds, consists of 16 Analog-to-Digital/Digital-to-Analog converters, and is capable of performing real-time Fast Fourier Transforms, linear predictive coding of speech and music, and recording, editing and mixing of input from microphones. As such, this system can operate in a real-time performance context, such as the Pierre Boulez composition, *Repons.* At the time of the above remarks, Moorer had already been at work for two years on the Soundroid, the Lucasfilm project which aimed to produce a digital audio processing station for the film industry.

For composers as well as the hybrid creature, the sound designer/artist, the benefits of an integrated environment with capabilities for synthesis, sampling

---

[1]Moorer, "A Conversation with James A. Moorer," Computer Music Journal
[2]Ibid.

and sequencing are clear. Although a traditional sound editor's job seems quite straightforward, this does not justify designing advanced, technologically intense tools with only the completion of simple tasks in mind. The Soundroid had far more power than most sound editors would even want to know about. The appeal of its design is that it didn't limit the user. The structure was open and nested, capable of expanding according the user's experience and imagination.

Soundroid pioneered a specialized nook of an industry that had started with two commercially available workstation designs, the Synclavier, designed by New England Digital Corporation, and a design called the Fairlight Computer Musical Instrument, designed by the Fairlight Corp. of Australia. These appeared on the market in the late 1970's. Both of these were initially designed with the composer in mind. They both featured some kind of dedicated processor, a customized keyboard (as this was pre-MIDI), an alpha-numeric keyboard and a monitor. The Fairlight in addition offered a light pen, allowing selections and some operations to be performed by touching appropriate areas on the screen. This included waveform drawing operations, which at the time was rather a novelty. Fairlight has recently gone out of business, and the Soundroid didn't make it fully to market. It is a volatile business. But despite the Soundroid's history as a prototype machine and not a commercial product, it encompassed so many interesting ideas that it is still instructive to go over its design.

Soundroid

The Soundroid project came out of the Droid Works research group at Lucasfilm. The project was initiated in 1980. It was designed as "a digital console, recorder, 'on board' effects unit, and total automation system[1]." This ambitious project encountered unfortunate marketplace difficulties, due partly to the expense of producing the system and the overall skepticism of the film industry towards computer-based tools, and partly from marketing difficulties produced by internal factors within the company.

---

[1]"Questions and Answers about Soundroid", Droid Works product announcement, 4/9/86

The Soundroid was conceived as a partner to the Editdroid, the computer-based random-access picture editing unit. These two units were envisioned as finally replacing the mechanical stand-by of the movie industry, the Movieola. Despite the closing of the Droid Works and the early retirement of the Droid systems, Soundroid remains an influential development in the area of workstations for the moving picture environment. It is also still in limited use at Lucasfilm, so I will speak of it in the present tense. The system incorporates advanced and enlightened hardware design by Andy Moorer and John Snell, and software that provides straightforward access to the capabilities of the hardware.

The system consists of two major components, the Audio Signal Processor, or ASP, and its interface, a fully digital console.

The Audio Signal Processor has been described in detail by Moorer[1]. The system was designed with many powerful features; since so many of its designers came from the computer music world a great deal of music processing and synthesis was embedded in the ASP. Moorer et al. felt that the major difficulties posed by audio for motion pictures centered around the large quantities of data throughput and simultaneous processing demands required to match existing practices.Matching this in the digital; domain would require large amounts of large amounts of numerical computation. In addition, there are bandwidth and updating requirements. The ASP addressed these issues by use of multiple data paths and complex architecture which allowed synchronous and asynchronous data exchange to occur at high rates without interfering with computation. Also, large blocks of program memory could be changed without affecting the sample data stream. The emphasis remained on sustaining real-time audio processing applications. Moorer wanted the system to address the the film sound editor's need to have large quantities of audio available with a minimum of fuss.

Briefly, the unit consisted of a controller and up to 8 digital signal processors. The unit is self-contained and semi-modular. The control computer is a 68000

---

[1]Moorer, "The Lucasfilm Audio Signal Processor," Computer Music Journal

with a Winchester drive, 1 megabyte of memory, and a high-resolution, bit-mapped graphic display screen. Each dsp was designed to handle 8 channels of audio at a sampling rate of 50 kHZ. With a maximum of eight dsp's per controller, the maximum channel number per ASP unit totals 64 times 300 megabyte disk packs provide storage, holding 42 minutes of monaural sound. Since this is not enough for a large scale film production, these disk packs are easily mountable, so that switching disk packs could at least be relatively painless.

The ASP was designed as a versatile engine. Theoretically, the ASP is capable of "generating large quantities of oscillators for use in frequency modulation synthesis, wave shaping, additive synthesis, VOSIM, band-limited pulse synthesis, discrete summation formulae, and other forms of synthesis[1]." Each signal processor can generate up to 60 oscillators, each with independent amplitude and frequency envelope. These processors could also generate many types of filters for use in equalization, reverberation, subtractive synthesis and noise reduction. Further capabilities include; phase shifters, harmonizers, linear prediction, Fast Fourier Transforms, compressors, limiters, spatial movers with Doppler shifting and pitch modifiers.

The console which interfaces to the audio signal processor consists of hardware knobs, high-quality Penny & Giles faders, switches, joysticks and touch sensitive strips. This in turn attaches to its own computer system: a central processing unit based on a 68000 microprocessor, memory, disk drive interface, graphics display (SUN high resolution, bit mapped display). This system was designed as a general purpose, software programmable interface, where the user could assign the hardware controls to control any aspect of the ASP. It was a virtual console which avoided hardwiring any of the hardware to definite parameters. It also was fully automated--any motion of any piece of the hardware interface was held in the console computer's memory.

This scheme elegantly answered to several different design issues. Does the system give the user access to the power of the system? Does it give her the

---

[1]Snell, "The Lucasfilm Real-time Console for Recording Studios and Performance of Computer Music," Computer Music Journal

information she needs to perform the job at hand, while restricting useless, irrelevant informations which is only distracting? The ASP was capable of so many functions that a literal mapping of function to control would have been overwhelming to a user. By choosing the controls and parameters herself the user could design an environment suited for her particular needs that day. Set-ups could be saved including patching and special effects. This strategy also allowed an open-endedness to the system by allowing the hardware to shift identities. It could function as a mixer or as a musical instrument. Since much of mixing, particularly in film, is like a real-time performance, the notion of respecting a 'feel' to the system supported both sets of possible users. The operator could set up a configuration and practice with it, adding to its complexity as she improved.

The console was initially configured to cover the tasks of film-oriented sound production, such as sound mixing, editing, equalization and specialized effects (fly-bys, Doppler shifts, limiters, etc.) But the broader goals of this system apparently were aimed toward music production as well, where these hardware elements could be user-configured into an instrumental context. Briefly, this interface was technically achieved by a signal path where flow of control data streams from the hardware knobs, faders, etc. to the 68000 microprocessor, then through the computer's Intel Multibus to the update system board, through the digital signal processor controller to the digital signal processors themselves. In turn, the displays could receive some or all of their control information from the console computer[1].

The software   followed a desktop metaphor, and favored icons and graphics, such as familiar buttons and faders. Although perhaps bordering on the cute, the screens I have seen (not many, I confess) are not overly busy and offer a good balance of necessary information without overwhelming the user with possibilities. A touch-sensitive screen completed the effort to make the screen a transparent interface.

The Soundroid incorporated features which seem far ahead of their time today. The design of the processor, which allowed for extensive potential in synthesis

[1]Ibid.

and signal processing, and thoughtful design of the virtual controller not only imbedded the practical possibilities for growth of both the system and the operator but invited exploration.

Current developments

Today, several kinds of digital audio workstation designs are available commercially, in addition to the Synclavier and AudioFrame. Most are weighted toward music production but a few have been designed with film and video post-production in mind. None achieve quite the ambitious level of Soundroid. I will mention just a few, to point out how the design schemes represent different strategies of design and marketplace.

Some manufacturers have approached the issue of user accessibility by trying to replicate familiar interfaces, as in the Opus system by Lexicon. They use a hardware emulation of a traditional mixing console as a front-end for a software-driven hard disk recorder. This system is clearly intended to replace the tape recorder and mixing desk, although it also includes editing capabilities in software. This system runs in the high five figures.

Another approach uses smaller-scale hardware designed to interface with small computers. The IMS Dyaxis system, for instance, consists of a set of proprietary converters and a modular hard disk recorder. This hardware is designed or use with a Macintosh computer, thus allowing the user access to other collections of software. The system comes with a software mixing program, MacMix, designed by Adrian Freed, but the system can be used with several popular inexpensive programs. This scheme attempts to offer an affordable package to low-end users especially those who already own a small computer and use it for sound-related activites. This system runs between the high four figures and low five figures.

In general the market divides between expensive hardware intensive designs and designs that try to balance the expense of necessary specialized hardware with existing small, affordable computers. The latter strategy can capitalize on the competition in the software market. The division centers around an approach that favors straightforward hardware and extensive software or customized hardware with less emphasis on software.

Most of these systems, regardless of size, now support SMPTE and even VITC time-code formats for video. This represents a hopeful trend, as it can save the user from purchasing an additional interface box. The model of the video edit decision list is also moving into sound effect editing software, and hopefully there will be more extensive use of intercompatibility with existing video edit decision list formats.

SPECIFICS ABOUT THE SYSTEMS

Conditions of the Evaluation

I received the AudioFrame at the end of August with the hopes of doing the majority of my sound production with it. Andrew Currie, an engineer with the company, arrived with the system and helped me install it and run through the set-up. At that time there was no manual, and I did not receive one until October. This was unfortunate. The system loaned to the Film/Video section was an early configuration of the system and had many software bugs. I also found that the sequencer was still in very primitive form, and this latter condition particularly made it seem unlikely that I would be able to carry out my plans with the soundtrack with the loaned configuration. I spent several weeks exploring the system and the possibilities which existed to undertake parts of the project. However, time constraints forced me to find the simplest procedure possible. I decided that it was important to keep all the work within a single environment, and I approached the New England Digital corporation for access to a Synclavier.

Ted Pine, of the New England Digital Corporation, expressed an interest in helping me with the project, and we started work in the New England Digital demonstration studio in September. This studio is outfitted with a Synclavier 9600 system and Direct-to-Disk hard disk recorder. There are a small number of external effects processors, such as the Lexicon PCM 70 digital delay. I should add that it was necessary to have an experienced operator on the project. The Synclavier is a big, complex system. One would need months to learn this system on one's own.

The AudioFrame

The AudioFrame is produced by the WaveFrame Corporation of Boulder, Colorado. The company is young, having started in 1986 with the express purpose of designing and manufacturing a unique digital audio workstation design. The system was announced at the October 1987 Audio Engineering Society show. It is important to emphasize that this system is going through an infancy at this time and undergoes constant revision. I was loaned for the purposes of this project one of the very early configurations of this system, and the company has in that time (since September, 1988) produced a substantial hardware update, many software changes, some of which are associated with the new hardware, others are corrections and extensions of existing functions. Some new software is on the horizon, notably a port-over of the sequencing program, Texture. I will examine these developments in more detail, but the above gives an idea of how quickly the system has been changing.

The AudioFrame system consists of two basic elements: the Digital Audio Rack, or DAR; and an IBM-compatible PC with a 386 processor. The latter runs MS-DOS, and its window environment provides access to the hardware in the rack. The AudioFrame is designed as a modular system (see illustration) The computer communicates to the DAR through a Token Ring Network. This network passes packet data at 4 megabits per second, and allows interfacing of both additional control computers and/or multiple Digital Audio Racks. The Token Ring handles all the protocols necessary to support a multiple user, resource sharing environments.

The rack is constructed as a frame which supports a synchronized bus. The user can easily slip individual cards into the bus, or remove them. The modular units are divided into the following:

Studio Control Processor or Controller Module

An individual control (SCP) module can drive up to four DAR's. This module is actually a communications computer, and it contains four independent MIDI ins/outs, SMPTE/LTC time code in/out, SMPTE/VITC in/out, house sync input, external sample clock output and metronome out. The unit also carries the connections to the token ring network. The board contains three 16-bit microprocessors and over a megabyte of memory, enabling it to sync to SMPTE/LTC while reading MIDI or driving other external MIDI devices.

Analog-to-digital module
This module is available in either stereo or 8-channel configurations. Analog audio is converted to 16-bit digital through this module at the rate of 44.1 kHz. From here, the digital audio data can be routed to any other point in the system. WaveFrame uses 18-bit converters to ensure that all the data maintains constant 16-bit quality. Connectors are user-configurable to balanced or unbalanced input.

Digital-to-analog module
This module provides analog output over 8 individual output slots. Four stereo pairs of AES compatible digital output are also provided, allowing work to continue in the digital domain for digital audio recording or PCM encoding. The outputs are self-calibrating, and, as with the modules above. This module also contains an on/off switch to disable dithering. Dithering can help maintain signal quality at low output levels.

Sampling synthesis module
This module contains 16 dynamically allocated voices, and carries two megabytes of random access memory on board, which provides 24 seconds of storage at a sampling rate of 44.1 kHz. This record time can be extended by lowering the sampling rate to 22.05, 11.025 or 5.5125 kHz, which would render record times of 48, 96, and 192 seconds, respectively. The hardware specification provided by WaveFrame includes a 32 million floating point operations per second digital signal processor, an internal sampling rate of over 23 MHz and 37-bit arithmetic. All of the voices operate in synchronization, which allows mixing to occur within the sampling hardware itself. For instance, this means that all sixteen voices could be mixed down through one D/A output.

14-Megabyte memory expansion module
This expands the random access memory of the sampler module. One expansion board results in an increase to 192 seconds of record time at 44.1 kHz, and each rack can accommodate two expansion boards. With two

expansion boards installed RAM grows to 30 megabytes, allowing over six minutes of monophonic recording at 44.1 kHz.

The Digital Audio Bus

The system revolves internally around the 64 channel digital audio bus. This is a switching matrix for digital audio (to be really precise, a time division multiplexed telephone-like switching bus) which allows all the audio information to remain in the digital domain as it travels through the various processes contained within the digital Audio rack. It functions like a 64 point digital patch bay. All 64 channels can operate simultaneously. The multiplexed bus structure enables any module in the system to communicate audio information every 350 nanoseconds at 24-bit precision.

This 'backplane' also carries the CommandLink network. Operating at 1 megabit per second, this network handles timing signals, SMPTE information, multiple MIDI streams and user commands from the computer. CommandLink's high speed helps avoid bottlenecks of data within the system. The network allows for the information equivalent of 32 MIDI cables, of 16 channels each. Practically this can be useful for integrating external MIDI data with internal AudioFrame processes. There are two MIDI inputs to the system which can allow input from external MIDI sequencers, as an example of this.

Any module can communicate to another in the system within 350 nanoseconds. 24 bit precision is maintained throughout, and all 64 channels can be operating at the same time. This allows the AudioFrame to maintain on-board mixing and effects processing capabilities without tying up the whole system. Practically, that means that a user should be able to send a voice through the DSP functions while another voice can be played with different processing.

The AudioFrame is not a synthesizer, it is a very high quality sampler. My loaned system consisted of one sampler, one 14 meg expansion board, one DSP board, one a-d converter board with two balanced inputs, one d-a board with 8 balanced outputs and AES compatible digital output, and the controller module which handles synchronization and external communications.

A single sampler configuration provides for 130 meg of storage on the hard drive of the PC, and 16 meg of ram memory--2 meg on the sampling board itself, and 14 on an expansion board. 130 meg supplies about 180 seconds of mono sampling or 90 seconds of stereo. Sounds can be recorded directly into the system through mike or line input, or can be uploaded through an everex tape drive. WaveFrame supplied a sound library on data cartridges for this model of their system. Sampled sounds can be manipulated through editing, looping and pitch transposition, or through the DSP functions of equalization and reverberation. This system does not currently provide for internal FM or additive synthesis, or modifications of samples on the level of harmonic structure. The system requires an external MIDI controller for real time musical input.

The system is modular. The design of the hardware originated with the company. This was a big part of WaveFrame's initial concept, and I think it is worth keeping in mind the extent of this undertaking when we consider the NED Synclavier and AudioFrame in relation to each other.

Software

Software for the system runs under MS-DOS and Microsoft Windows. WaveFrame divides the functions of the system into three major software sub-divisions, SoundProcessor, EventProcessor and MIDICad. The first two divisions are fairly self-explanatory. The latter area covers the DSP programming. This is where mixer and reverb set-ups are stored, as well as a tool kit for development of user-defined control setups. I will briefly discuss the software here, and refer to the reader to the AudioFrame manual for further details.

SoundProcessor

As an overview, SoundProcessor contains windows for such tasks as defining sampler setups, recording new samples and performing waveform-based editing and looping functions, MIDI and DAC output assignments, keyboard setups, and creating "instruments". An instrument can be comprised of one sample, or signal, or several. (see Fig.) The SoundProcessor application is the most complex and detailed of the three. This is the center from which sound elements are directed internally throughout the system and externally to MIDI devices.

There are some nice features in the editing application. The crossfades sound smooth, and they have approached autolooping in a clever way. The software looks at patterns of zero crossings and tries to find the nearest identical patterns. It then offers some possibilities to preview. There is a scrub feature which allows the emulation of cueing sound over a tape head--this is assignable to a controller such as a keyboard pitch modulation wheel. There is also a *trimbin* window which collects outs of edited material. I encountered many bugs in this area with my machine, to the point where it became quite difficult for me to evaluate the functionality of all these features. For instance, it was difficult to determine whether it was possible to undo edits, or edit easily between windows/individual samples. These problems are being eliminated, and I received a software release which did show improvements, but with the reliable and powerful sample editing programs available in competition it behooves WaveFrame to foolproof this area. After all, sampling is the system's current main function, and the editing interface must become truly dependable to warrant the expense of this system over other alternatives.

As mentioned earlier, the AudioFrame does not at present offer any internal synthesis capabilities. The company seems headed toward developing additive synthesis in the future, and I understand that development is already underway in that area[1]. The possibilities for sound modeling at present are situated within the vocabulary of digital sampling. This means modifications through editing, looping and crossfading loops; reversing samples; pitch transposition; addition of vibrato through Low Frequency Oscillators; and modeling the amplitude envelope through adjustment of attack, sustain, and release parameters. I found this part of the Instrument modelling difficult to use. This is due in part to the numerical orientation and to my personal bias toward a graphic interface in such tasks as amplitude envelopes and even keyboard assignments; it is also due in large part to the gestural editing feature that controls these parameters.

[1] Meyer, "The AudioFrame explained, part Four", Music and Technology

Gestural editing deserves a mention on its own. This term refers to a feature that allows the user to change values or move things on the screen, such as knobs, by holding down one or two of the mouse buttons and sliding the mouse until the desired position or value is found. This feature is quite active in the Instrument window and in the mixer/reverb applications, where knobs and faders can be grabbed and adjusted by moving the mouse about. Initially, and somewhat abstractly, I was drawn to this idea because it seemed like a way to introduce the notion of "feel" into a software environment. My practical experience has led me to temper this enthusiasm. The execution here, particularly in the numerical applications, was very difficult to control. This scheme might last in the knob/fader applications, where the target objects are visually larger (See MIDIcad description below).

I would like to see sound modeling become more powerful. This may occur naturally as the interface improves for instance having graphic amplitude envelope shaping will make that feature more accessible. The more dramatic effects can be achieved by running signals into the reverb and re-sampling them, but until one sorts out the signal paths this can be very cumbersome. I would like to see filtering and other effects accessible through SoundProcessor. Eventually, perhaps quite soon, I suspect that the separation between the power of the signal processing and SoundProcessor will become quite problematic--even though the cad applications will remember setups. There is a certain amount of processing I would want to perform on the signals themselves, without having to get involved with exiting SoundProcessor in order to call up a a preset. Re-sampling, although an important feature (and important to be able to easily execute) does not seem like the sole solution here. This seems like a situation where the ability to communicate remotely, to call routines, to and from windows could be powerful.

SoundProcessor has some information delivery problems. The Instrument window contains many possible parameter adjustments, such as LFO on/off, LFO frequency, etc. But they are placed in grids, and since these adjustments occur as numerical values this can result in a very busy screen if the Instrument has many signal components. (fig.) This is a difficult area of applications to balance because there are so many different kinds of parameter adjustments this application is trying to allow for. These range from details about sound parameters to keyboard layouts and routing instructions. I think that some of these tasks need more visual aids, such as routing (see "Mental Models", below), whereas others could be simplified by replacing the numbers with graphic forms.

EventProcessor

EventProcessor contains the sequencer, the 'transport' window which serves as a controller for the sequencer or external synchronization and a separate edit decision list window. Synchronization controls are located as options in EventProcessor. The AudioFrame will sync to both SMPTE time-code or VITC, at 24, 25 or 30 fps. It recognizes both drop and non-drop frame time code. Additionally, this system will generate SMPTE, enabling striping of video or audio tape. It will also sync to other MIDI sequencers.

The sequencer is very primitive and stands out as the weak link in the AudioFrame's current functionality. It offers no editing features and only takes performance data input. It also has no link to the edit decision list, which is the other sequencing application. A new sequencer is on the way. This will be a port-over of an existing program called Texture, which Roger Powell designed for IBM and IBM-compatible machines. I have seen a beta version of this software, although I have not had an opportunity to use it myself. It clearly presents an improved sequencing environment, allowing global views of track data and local views of note and MIDI data. Hopefully, it will be possible to perform global edits across tracks--this is a useful feature for film and video appplications because of the parallel to shot lengths.

The edit decision list mirrors a video-style time-code based format. Events can be entered at specific time-code addresses, and events can be defined as either single notes or sound effects. A good feature here is pitch transposition and volume adjustment from the edl itself. Here is an example where the window offers important controls--an independent set of important instructions--which can be executed from that environment without having to exit. In this regard it strikes me as unfortunate that sequence on/off instructions cannot be listed as edl events. This reflects some crucial communication problems which I will examine further below but suggest here that for the purposes of film and video applications it really is crucial that all sequencing features should have a mirror in the time-code domain. That is, a sequence should be viewable with a reference to SMPTE time and associated edl events, and the edl environment should be able to support, or at least reference, sequences associated with shared time-code addresses. Otherwise, a music and effect editor working separately might not realize that they are both loading a certain portion of the track with events. As a user who wishes to work simultaneously in the domain of "effects", ambience and music, I found the division of EventProcessor into discrete functionalities very inhibiting creatively.

MIDIcad

This application presents the most exciting direction of WaveFrame's design strategy. These software tools access the DSP board, and currently include the mixer and reverberation setups. However, there are development tools here for user-configurable controls for this board. This area presents the closest thing to a virtual controller in the Soundroid model in either the WaveFrame or New England Digital current design scheme.

Some of these setups, or what WaveFrame calls 'views', are available already configured, such as mixers with different i/o distributions and equalization settings. Features of the software mixer include parametric equalization, high and low frequency shelving, phase adjustment for matching phase, etc. One can save views and also presets, which refers to the settings of controls rather than the entire layout.

The reverb comes with several interesting presets from WaveFrame and sounds very good. It can also be configured by the user. This feature represents an effective use of the DSP capabilites of the system.

Soundstore

Soundstore refers to a major hardware and software update that WaveFrame released not long after I received the machine for evaluation. The hardware addition consists of a SCSI adaptor board, at least one high-capacity hard drive, and a streaming tape backup unit. The latter take 8-mm video-tape. All fit into the DAR. Additional SCSI adaptors cn be added, and up to four hard drives can be installed in one rack. This setup can provide from 90 to 900 megabytes of storage.The main improvement here is in the time it takes the system to up- and down-load sample RAM from disk. This configuration does not allow playback of sounds longer than what can fit into RAM, however. It does not allow playback of large sections of continuous sound.

Soundstore comes with a file mangement system with an associative database. This represents a marked improvement over the previous DOS file structure for locating sounds. It is also possible to audition sounds while searching through the database, rather than having to load them into RAM in order to hear them, a procedure required by the non-Soundstore machines. This was highly inconvenient, as load time was considerably slower to begin with in the older models.

A metaphor for the overall software design

In general, the AudioFrame software seems very problematic. It can be idiosyncratic, self-referential, illogical. There is not enough feedback in the system as it currently stands. For instance, it is not always obvious when the system is working on a task, and saving is an operation that must be performed frequently in each different window environment.

I would describe this as a 'discrete room' approach, where different functional units are represented as windows, or rooms. Within the rooms the user can find various tools, but to access others, she may have to get up and go into the next room, rummage around, and finish a task in there before she returns to the original site of operations. Thus, to record a new sample for an instrument window, the user has to open an existing or new Instrument from the SoundProcessor menu. Then, from the Instrument window she pulls down another menu to find "record new stereo/mono"...

The advantage to this Windowing environment is that one can get the various work areas up and somewhat available on the screen for reference or further activity. I say "somewhat available" because this advantage becomes reduced by additional visual clutter, the tendency of screens to hide behind others and elude the mouse and the operator, and slower screen drawing operations which can make the re-draw of just moved screens very frustrating. In my system this was compounded by some bugs which resulted in crashes if I tried to execute something while a window was still re-drawing somewhere on the screen.

I ultimately developed mixed feelings about the Windowing environment. I felt that I spent a lot of time opening, closing or moving windows around to their optimal position. If I had all the windows open that I might need for reference or feedback, I began to loose track of them. Sometimes I felt that the problem resided simply in working out the correct balance of types of information in each window and correct presentation. Some windows seemed to have too much information, such as the Instrument window, where there exist many numerically-based adjustments for different parameters and some not enough, as in the routing mapping.

Windows present what their name implies, access to a sphere of activities. But it would be nice to be able to reach through windows sometimes if we need something on the other side. In the EventProcessor, the transport window will talk to the sequencer and the edit decision list windows, but the sequencer can't talk to the edl. Nor can the edl store a instruction to the sequencer. I have mentioned this as a practical problem but it also reflects a theoretical one, because any way one divides up tasks in order to fit them into windows is bound to be a bit arbitrary. There are no real standards to measure where the divisions should go. A user, such as myself, who has firm ideas about the ability to execute one task from the functional environment of another will be very frustrated and disappointed by the interface if it does not fit her mental model of how these tasks relate. I have that notion because my experience has informed me that these tasks are somehow related.

Overall Evaluation

I was immediately struck by the apparent simplicity of the AudioFrame's modular layout. I admit to a bias towards the idea of modularity. It allows the buyer to configure the system according to her needs and budget, which seems to me both practical and accessible. Also, the AudioFrame is modular in a simple enough manner that its owners can perform most of the modifications themselves, without having to call in costly company reinforcements.

The system is small for its capabilities and this adds some flexibility for situations where space is a premium, and for site or location work. It would not, for instance, be difficult to set a system up quickly in a theater for live music or effects work.

This system is more affordable than the Synclavier. The pricing scheme ranges from about $40,000 to $90,000, with individual boards running between $2500 and $10,000. The Synclavier can run between $44,000 and $109,000, with the Direct-to-Disk adding an additional $58,000 for a four-track configuration at 25 minutes of track time,, to $134,000 for a 16-track, also at 25 minutes.

Most of my evaluations are based on the earlier model of the system, as this was the machine available to me. However, I did see the new hardware update, called SoundStore ™, and a beta version of the new sequencing software, Texture ™, which was designed by Roger Powell initially for IBM-compatible machines and which she has been porting over to the AudioFrame environment. I will start this evaluation by mentioning broad-based issues.

Mental Model

This is quite literally a black box system. With so much of the activity taking place being physically hidden to the user, it becomes very easy to feel lost. It is particularly easy to get confused in routing signals around the system. It is wonderful that there exists a 64-channel digital patch bay inside the system, but it is not wonderful if the user cannot figure out where signals went, why there is no sound coming out of the system, etc. In such as situation it becomes imperative that software provide feedback, information and illustration. Software carries the responsibility to make the system visible. I think a visual representation of the patchbay, for instance, would be very useful. Or perhaps some kind of map that would allow the user to trace the path of the routing she is attempting, in order to see where she might have erred.

In general I find that the software does not adequately inform the user about the status of the machine. There are times when the machine is working, for instance, normalizing a signal, and it is not clear that it is active. This results in unnecessary repetitions of actions. The system should always inform the user about potentially destructive operations, (one should not have to read the manual!) which does not always occur--for instance, cross-fade looping is a destructive process which does not announce itself.

I see that the challenge for WaveFrame lies in bringing the software to a point where it really does allow the user access to the power of the system. The software tends right now to be chaotic and idiosyncratic. It does not necessarily follow a logical path: the user has to remember many details about which pull-down menus contain which commands, and which larger windows they reside in. The amount and kind of commands hidden in pull-downs throughout different windows in the system is quite varied. This makes the execution of certain tasks window/location specific, which can result in a bit of hunting and longer learning period.

DSP pros and cons

By carrying on board digital signal processing the design scheme promotes a fully integrated system that theoretically reduces the need for many peripheral devices, such as external effects processors and mixing boards. They see the flexibility inherent in such a structure as a design goal, which I think is admirable. However, it is not obvious to me that the internal software mixer provides an adequate replacement to a hardware mixing board. The software buttons are very fussy to use, the screen is cluttered, and the ergonomics of poking with a mouse, even though they incorporate some gestural editing features here do not provide an equivalent tactile sense.

But overall, the potential of this DSP board seems like the most powerful and unique feature of the AudioFrame. The quality of the reverb is very good, and if the company should decide to develop the potential of this board further into synthesis I imagine other applications will sound equally good. The power here seems underexploited, since the board has already been through hardware development. Hopefully, the company will decide to explore this direction.

SYNCLAVIER

It is important to establish that this system has been around longer than any other existing computer-based audio workstation. It started out as a project at Dartmouth in 1977 to build an interactive computer-based instructional tool for teaching music theory and composition.

At that time the focus was on digital synthesis and sequencing. Synthesis at that time was still based in analog technology and principles, so the first innovation for the system was its focus on digital protocols for all operations. From a very early stage the designers of the Synclavier aimed to incorporate real-time composition, recording and performance with the capabilities of a synthesizer. The Synclavier remains capable of synthesis as well as sampling, of sound analysis and intensive sequencing. They have not sacrificed functions to newer innovations, but have tried to integrate the new features. The result is a powerful array of possibilities, although somewhat sprawling and perhaps intimidating to a user. Here it suffices to remark that the aims for the system overall have always been large, and the system physically reflects that. It is very hardware intensive.

The original design sat on Dartmouth's mainframe computer. But it became necessary to design specialized hardware and a dedicated minicomputer in order to deal with all the specialized needs for computations on the audio materials. At this point much of the system is manufactured by NED. The perhaps sole exception here is in the various drives used for back up and storage. NED now offers three different major configurations of the system, the 9600, 3200, and Post-Pro series.

NED has also come out with a very powerful Direct-to Disk hard disk recorder which is integrated into their top-of-the-line systems and available as an option in the 3200 configuration. The D-D offers continuous recording capability of 44.1kHz digital sound at over 28 minutes per track. Each 'track' is actually a 300 Mb Winchester hard drive. The module is available as a 4-, 8-, 12-, or 16-track recorder. Sounds are digitized to hard disk at the same sample rate as the Synclavier --16-bit, and the unit is fully integrated into the system as a whole. That means that control of the hard disk recorder is possible from the keyboard, and all the sequencing and editing software for the recorder appears in the context (next door to) of Synclavier software. By sharing software in this manner live material can be edited or manipulated simultaneously with sequenced material.

The operating environment of the Synclavier is distinct from the recorder but connected. This would allow the following situation; recording live material into the D-D, such as vocal or acoustic instrument performances, and being able to view sequences and samples within the Synclavier environment simultaneously with the live data. For filmmakers, it would allow location sound to be dumped into the hard disk and synchronized to picture while composing sequences from timbres available in the Synclavier realm. One could also sample from that live sound on the hard disk recorder into the Synclavier environment and put it up on the keyboard as an instrument. I will get into this more deeply, but before these flights of fancy can make much sense, it's necessary first to make a more methodical outline of the system.

For the sake of simplicity in the following description I am looking at the various components of the system in the context of two categories, custom interface hardware and computing/storage resources.

The most immediately obvious piece of proprietary hardware is the keyboard. This has a bank of buttons, a data display, and a data entry wheel for specifying exact values in such operations as chorus depth/delay and attack/decay modifications. This is really a command interface for the system as a whole, allowing many tasks to be completely executed from the keyboard. Some of these include:

Initiating recording and playback:
into the 32 track recorder/sequencer, also includes external synchronization and SMPTE mode. The features of the sequencer here include looping, segment chaining, transposition, track justification, track volume, tuning. The recorder provides the expected features: rewind, fast forward, punch in, erase/record, continue.

Assigning tracks

Adjustment of timbre parameters, including partial timbre select, which allows timbre mixing on one key or regions of the keyboard. Also includes access to FM synthesis and control of partial volume, partial tuning and harmonic select.

Keyboard control; includes harmonic and decay adjustments, chorus, arpeggiate on/off and rate.

Timbre control; includes vibrato depth and stereo control.

Real time effects; assignment of effects to controllers such as pedals and modulation wheels.

The keyboard controller is a fairly significant piece of equipment, larger and heavier than most other commercially available midi keyboards. It has a reassuringly solid feel to it, but this kind of equipment demands space. It's the first intimation of the scale of the system.

The controller uses rs422 protocol to talk to its dedicated cpu. In a strict technical sense, this cpu *is* the Synclavier. It boots from a separate Winchester drive. The cpu can be addressed by other controllers, such as a separate MIDI device, a digital guitar interface, a remote controller/editor/locater, or simply from the Mac ll workstation. The architecture branches out from this central processor. (see illustrations) The FM synthesis feature and DSP capabilities are not built in, but exist as separate units. Thus, the modularity of the system exists more in the model of hardware and rack mounted units than in one of swapping cards.

NED prides itself on the quality of the audio coming out of their environment. They offer variable sampling rates up to 100kHz. They designed a proprietary anti-aliasing filter to handle any phasing artifacts. At the output stage they have combined each 16-bit DAC with a separate 12-bit amplitude control DAC on a per channel basis. The idea here is that the decay portion of sampled sounds will sound "as clean and present as the attack."

Software is structured as separate pages per application within an overall structure of catalog and sub-catalog directories.

The Synclavier as a tool

The Synclavier is more than a sampler, or digital tape recorder. It does offer synthesis, analysis and re-synthesis. Further sound modeling can occur by programming partial timbres on the keyboard and adjusting the amplitude envelope. The keyboard becomes a useful tool here, allowing parameter changes to be correlated with pressure and after-touch.

Synthesis occurs as additive synthesis. A basic waveform can be constructed by adjusting the amplitude of 24 harmonics. The harmonic structure can then be modified further by applying frequency modulation. Frequency modulation is controlled by buttons and the value dial on the front panel of the keyboard. This sets the frequency of a modulating wave or selects a fixed modulator frequency depending on whether the value is positive or negative. The depth of modulation is controlled by specifying an harmonic envelope. This scheme does not allow for the kinds of complex modulation relationships that the Yamaha model does. But the carriers in the NED model tend to be more complex. The interface is easy to use, as well. The controls for determining the harmonic envelope are also set through the hardware keyboard interface.

Resynthesis allows the analysis of sounds, the determination of their harmonic components, and the re-construction of that harmonic structure. The re-synthesized sound is broken down into segments called timbre frames, each of which consists of a set of harmonics, a delay time, and a splice time and shape and a peak volume level. In addition, each timbre frame can have a pitch envelope, enabling pitch fluctuations to be programmed into a partial timbre. Timbre frames can be manipulated as single units or spliced together. A re-synthesized sound on the Synclavier can be modified through frequency modulation, or through modifying the harmonic structure. Re-synthesis is a non-real time activity in the Synclavier, and my limited experience with it suggests that, like any fine level process, it takes some time to learn to use this to best advantage. But this does offer an intermediary stage of sonic manipulation between sampled sounds and purely synthesized timbres, which the AudioFrame does not provide for at this stage of its development.

Sampled sounds may not be modulated through FM, nor can the harmonic structure be manipulated. However, the ability to so easily layer partial timbres on the keyboard, in addition to various real-time effects also achieved through keyboard interface, offer many possibilities for blending different types of timbres. For instance, out of four total, one timbre could derive from a sample, another as an FM, another as a re-synthesized sound, etc.

It is clear that the Synclavier offers some powerful methods for sound modeling. This is an important direction for the workstation to explore further. NED is currently re-writing the re-synthesis application and it will be interesting to see if anything new comes out that. The advantage of building custom processors and hardware should be in the enhanced sonic palette made available to the user. This seems to be a crucial area to exploit if an expensive system is going to try to prove itself against the increasingly powerful, small systems. As DSP chips, like the Motorola 56000, become cheaper, we will see more small custom boards like the DigiDesign Sound Accelerator available. This board slips into any Macintosh with two megabytes of RAM and is configurable to allow equalization, stereo effects and conversion, all for between $1,000 and $4,000. The AudioFrame seems currently more active in the real-time digital effects sphere, yet remains in the sampler domain. NED has not as fully explored the possibilities of real time digital filtering and effects, though they have announced plans to develop more DSP based on the 56000 chip. One feature which is eminent is time compression and expansion, which would operate at 50-100% of the original value and be applicable to either individual events or completed mixes.

Evaluation

The Synclavier is a powerful but somewhat sprawling system. It is complex to learn, but it seems that after that initial period operating can proceed quite quickly. It has had years of development during which many of the bugs could be determined and obliterated. NED did not stop developing software, which seems like a terribly important momentum to maintain for these companies. Improving software or hardware interfaces to the render these systems more accessible seems key to survival of this inclusive design approach. NED also benefited from having some early development occur in an academic environment, without the pressure of bringing a product to market.

Because of this constant development NED has reached a fairly refined stage of design in their software. They have managed to move beyond many of the pitfalls that still plague AudioFrame software, although the arrangement and number of application 'pages' can be initially confusing and overwhelming. The screens also tend to be graphically busy, but at least provide generally good visual feedback. The design allows the user access to most of the information needed for each application page on that page, even if it does promote some redundancy through the system as a whole. They are moving toward creating more user-defined macros so that keystrokes can take over from the mouse. As far as saving, the system updates itself and reliably sends error messages about destructive operations. Editing applications for sequences and sample operations include an undo command, which seems obvious but is not always thoughtfully provided. NED seems to have these seemingly small but important and time-saving details well in hand, and this means that operating can proceed smoothly, with the focus on creativity and not fixing the machine. This is not to say that the system doesn't ever crash, or mysterious problems do not occur, for they do. The system crashed inexplicably a few times during my project.

As an example, several sequencing format options are available, depending on whether the overall context requires a time-code reference (SMPTE or VITC), MIDI data or note display. It is very convenient to see a sequence in these different formats. One can add MIDI events or "cue" events--the latter could be sound effects, voice-over, etc--all of these are under word processing control. It is possible to backtime events and ripple edit--in other words, the sequencing software for time-code related projects incorporates the specialized needs of motion picture operations but does not rule out those of the composer. They are integrated, not divided. I would like to see more of this kind of integration of specialized applications, as opposed to dividing the labor. NED is preparing a configuration of the Synclavier, the Post-Pro ™ specifically for the post-production market, and I hope that they do not start carving up this versatility too much.

Each list of events or groups can be assigned individual volumes (up to 200 such lists are available) of output routings. I think individual track volume control is very useful in the composing stage, even if ultimately a more extensive mix is planned.

The Synclavier design features seperate RAM for timbre storage and sequencer memory. Both of these can be configured by adding cards which range from $1500 to $5000 per 4 megabyte card. Dedicated RAM allows flexibility in designing sounds, as it becomes easier to have several kinds of sounds in RAM simultaneously without having to worry about memory management. This is a nice feature if one is working with long samples.

The addition of the Direct-to Disk recorder allows global editing to occur as well, in the from of cutting and pasting tracks. This finally allows some real work with continuous sound elements, instead of always having to chunk them down to sample size pieces. Sound elements can be swapped back and forth in digital domain from the Direct-to-Disk and Synclavier. For instance, one can move any live sound, vocal, instrumental or ambient, from the hard disk recorder to the Synclavier to edit it, or do a pitch correction, then send it back in sync to the Direct-to-Disk. This kind of invisible swapping allows great flexibility. The on-line optical WORM drive makes it possible to bring sound effects or timbres stored in the optical library to either the hard disk recorder or to the Synclavier to be modified and played from the keyboard as a sequence. This is really useful, as each platter can hold up to five and a half hours of sound, but the platters are $300 dollars apiece and the drive itself is $29,000.

The hard disk recorder is an important addition for motion picture-based activities. It allows synchronized lock-up to picture in under a second, no machine search time, and precise punches in and out without the slippage which can occur with sprocketed media. Any event can be slipped. For serious projects in the motion picture area access to large continuous audio segments must be maintained. Moorer caught on to this very quickly in the ASP design for the Soundroid. WaveFrame was planning to release a hard disk recorder as close as possible to the release of the Digital Audio Rack, but couldn't manage it at that time. It is very problematic for a workstation design currently on the market to lack a hard disk recorder-type feature, and I believe this will arrive in the AudioFrame environment soon.

The NED design does bring up a consideration in hard disk recorder layout and management. In their scheme there is a direct mapping of disks to tracks, that is, if a user thought she might have need occasionally of a 16-track recorder, she would have to buy a 16 disk configuration of the Direct-to-Disk in order to buy that number of separate outputs. In the NED model the user must commit sounds to specific disks/tracks. Memory allocation may result in storing silence as well, if that is seen as being part of the overall project, or track segment. In this sense it emulates the model of a tape recorder, where space on tracks can be freed up by bouncing. Here, sound is dubbed in its entirety to other tracks. Because of the dedicated disk/track relationship this situation also occurs with the Direct-to-Disk-- one cannot mix and match sounds and source. The flexible solution here would be to allow the user flexibility in allocating channels or track time--that is, to have 30 minutes of stereo coming off the disks or a reduced number of minutes which run through more outputs. The NED approach does emulate a familiar model, and it does work, but it is expensive on a per track per second of sound basis. It presents an unfavorable situation to the less well-heeled user, who cannot afford a large number of tracks. Upgrading from a four to an eight track, for instance, costs $24,000, with the overall recording time still constant at under a half-hour. To increase overall record time represents another cost increase; 25 additioanl minutes of recording time on a 4-track runs $17,000, and on a 16-track, $ 64,000. So, it is not a trivial cvonern to reconfigure the Direxct-to-Disk system. It does work, however. The quality is very high, and the system is available now. For certain users, such as big recording studios or broadcast television, these considerations remain the determining factors. But the expense of such systems puts an important tool out of the reach of a huge segment of a creative population, thereby widening the rift between commercial and independant productions.

The extended growth of the Synclavier system has led to some problems. The somewhat sprawling nature of the system physically is one result, as well as the number of pages of software which contribute to the time it takes to learn and comprehend the system. The Synclavier presents a different version of the mental model problem we discussed in relation to the AudioFrame. Here, the challenge is to present a simple, overarching presentation of the system, a distillation of the complexity. The manuals for the system run into volumes.

For all the difficulties I eventually had with the AudioFrame, I learned to use it in a basic way very quickly. I could sample sounds, and do simple sequences almost immediately. I doubt that I could have executed those tasks as quickly on the Synclavier, had I been on my own and armed with the appropriate manuals.

The conversion to the Macintosh ll as the computer control has improved the graphic dimension and provided a much better interface environment than the previously used DEC displays. I believe that at the time that these displays came bundled with the system the software was numerically based, which would have been quite difficult to use. In addition to gaining a bigger display area and higher graphic resolution, the switch to the Macintosh allows software flexibility, since through MultiFinder applications other than the NED software can run on the host.

It was unfortunate that timing in WaveFrame's development schedule prevented them from taking advantage of the Macintosh ll as a front end. Providing users with access to some of the currently available Macintosh software is a plus.

The Synclavier is a major investment in space and financial resources. It is also presented as a modular system but it is not obvious to me that one would want to downscale this system too much. As discussed above with the hard disk recorder, potential problems are avoided here by spending more money.

For instance, the keyboard adds some expense to the system, and NED now offers a stripped-down version where the keyboard is graphically emulated on the Mac ll, which has become the front end for the system. From the promotional illustrations I have seen, this emulation is distressingly faithful. I have not used it, but to look at, the replication of all the hardware buttons as software ones makes the screen terribly busy. A user would spend a lot of time poking at them with the mouse. The shortcuts offered by this hardware configuration are negated in a software interpretation that remains so faithful to its parent, three-dimensional object. In this case, saving the $10, 000 cost of the keyboard may not be worth it.

Storage media

With any large-scale project which uses the Direct-to-Disk recorder, one has to do extensive back ups across several different kinds of storage media. These include: floppy disks, Kennedy tape cartridges for backing up the Winchester drives, high-speed digital tape back up cartridges for the Direct-to-Disk recorder itself, and the optical platter for extensive use of sound effects. This creates archival and record keeping complications, as well as adding an expense to productions.

Marketplace Considerations

What should the ideal digital audio workstation for video and film look like? Should it answer to the needs of the industry, and its labor structure, to the detriment of other users, probably independents? People who do not, for whatever reason, operate under the same financial structure as the industry will probably want a more versatile machine. For those who make the investment to either buy a system or pay for access on a limited budget, the ability to address several stages of the process in one environment will make the expenditure worthwhile--for instance, mixing and mastering. Collapsing expensive stages of post-production into one expensive stage(!) may actually make some sense in that scenario. Artists and independent film and video people will probably embrace this kind of technology more readily than the average union post-production person. The former are always happy to see tools appear that give them more power and flexibility. But the union is built on the division of labor.

The structure of the unionized film industry, represents a conservative force that could well keep this technology at a very primitive or reduced state of functionality if they become the arbiters and targeted market. This would be very unfortunate.

For a manufacturer, however, these differences, the working practices, and the attitudes associated with them must present yet one more confusing aspect in a sea of questions around user identity. This remains even more problematic as these systems still tend to be very expensive, more expensive than most of the people most likely to be interested in owning one could afford. For instance, one of these would probably be enthusiastically received at media access centers, where the client pool is very mixed, and people are working with limited budgets. Independent productions usually function on small budgets, so a system which only requires one operator yet can handle all the stages of audio post-production (and retaining good quality) would be very appealing. Independent productions and media access centers are also without unions and the subsequent restrictions. But that enthusiasm would be contingent on the cost both to the center and the client.

This predicament is easier for New England Digital than for a new company like WaveFrame. New England Digital just has to scale down the Synclavier, since it was developed as such an all-encompassing system over so much time. They can carve it up and present different packages, as they already have done by offering the "Post-Pro" (a package customized for the post-production industry) and a scaled down Synclavier system, the 3200. A new company may have to make these potentially limiting decisions at the design stage.

I place myself firmly in the camp that would like a digital audio workstation to be as powerful as possible. Of course, I would also like to be able to afford one. So, in the long run, although the power of the Synclavier remains very seductive to me, the modular approach of WaveFrame is one that I hope can also survive and develop in the industry.

It is highly problematic that these systems remain so expensive. There is a tremendous amount of development activity in the software domain for small computers which drive MIDI devices. Systems such as Dyaxis offer good quality converters and hard disk recording at comparable sampling rates for much less money. This modular approach of mixing and matching smaller scale products has some real benefits: the competition in the software market particularly is keen enough that revisions and improvements occur there very quickly. And, as mentioned earlier, small converter and DSP boards for insertion into small computers are now becoming viable, inexpensive. and available. All of these developments indicate that the smaller-scale producers will be seriously challenging these high end systems. This puts the pressure on companies such as WaveFrame and New England Digital to keep developing good software, and not get too complacent.

Personal musings

This experience has not been without its frustrations. It is difficult and somewhat crippling to work with one of these systems before it is mature, as as I learned with the AudioFrame. The danger of having a single software and hardware environment is that if it is not well worked out, the user becomes truly stuck. Just a few limitations, such as, in my case, the primitive state of the sequencer, can result in operator despair. A user who has spent a large sum of money to acquire an integrated workstation will not want to rig a makeshift solution by, for instance, sequencing the system through MIDI from another program, or another machine. Nor should she have to.

The Synclavier is still subject to occasional inscrutable lockups and other odd behaviors. The system is complex enough that even an experienced operator can become lost in the system. It clearly takes many long hours to master this machine.

The state of the art is still very young. But overall, and despite the occasional frustrations which arise from the exploration of a new knowledge base and subsequent adaptation of personal methodology, I remain convinced that these tools represent a great improvement in functionality and creative expression for the film or video-maker. Having explored this new territory, it is difficult to imagine going back to the old practices.

CONCLUSION

As a result of these thoughts I have come to hypothesize a kind of advanced language of image and aural processing as a new level of psychological exploration in new media. That is, I began to consider what movies might look like if we could apply equivalent tools and aesthetics found in computer music to the process of image-making. The tools for such an activity seem close. Here, I am thinking of some projects currently under development at the Media Lab, specifically, OBVIUS, or Object-Based Vision and Image Understanding System, developed by the Vision Science Group. This system uses the object-oriented programming environment of LISP to make images into manipulatable objects, and theoretically could integrate sound objects as well.

The benefits of applying a computer-based environment to enlarge the sonic dimension of moving pictures seem great enough now to warrant further explorations by filmmakers. These benefits probably remain strongest in the transformational and and sequencing applications. The generative power of the computer--its ability to define the parameters of a composition is, I believe problematic in the context of developing a language for filmmaking. In the long term I see some very interesting possibilities for an algorithmic approach to placing pictures and sounds, but it is not obvious that pure indeterminism in this area would produce interesting results. AI would help here: algorithms that had intelligence about the set of sounds and pictures at their disposal, for instance. But random generation of possibilities does not seem to capitalize on the potential of the computer. For the moment it seems that the asset of computer in the realm of motion picture expressiveness lies in the the direction of control, not the abandonment of it. That is, the precise manipulation in the sonic domain which can occur either as timbral or sequence articulations, in combination with the value of the images conceived in relation to these compositions.

The tools to trap time are going through another stage of evolution, from the chemical and magnetic to the digital, placing the artists who use them in between old assumptions and new questions. In the realm of image-making, for instance, fixing the image digitally lacks the standard of objectivity which has so informed photography. A photographer controls the silver halide molecules in her emulsion globally--he cannot operate on the level of each of those molecules. These constraints have led us to establish certain aesthetic and critical standards about photography and cinematography. One of these concerns authenticity and the poetry of the suspended moment. This depends absolutely on the absence of artifice within the process of producing the image. The process, by its inherent plasticity, opens itself up to doubt. What is lost here? The resonance of the lived moment--lived by the photographer as well as her subjects?

But as we have gained access to more of the material fixed by time we have been able to more deeply explore our experience of time, and use this creatively. Thus, the resonance comes to reside not in some kind of authenticity of the image itself, but from an authenticity within the perceptual experience of the image. The emphasis shifts from the object itself to the act of perceiving the object, as a cognitive event. Our questions begin to concern form in sound and picture objects as they are represented internally, in the mind, and communicate to other minds, with a focus less on the external, superficial value of the objects.

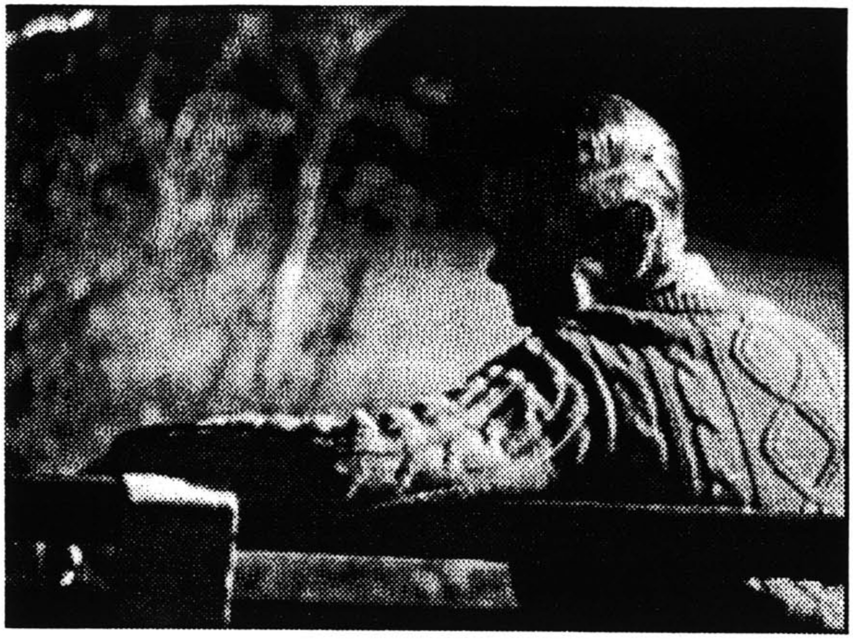
Digital technology allows us even greater access to the constituents of time--the grains become more discrete, as pixels on a screen with red, green blue values, and also more controllable. Control is what we buy, or think we buy, through digital tools. Control of material, control of time. We must re-evaluate the notion of time in relation to our current ability to pin it like a butterfly and dismember its various features.

We must consider what we are trying to express to people through our manipulations, because ultimately we are manipulating their time, as well. What are we going to offer? A new tool is valuable insofar as it allows the user to execute something new, or an old task better. And if a piece of work is solely interesting because of the tools used in its creation, than what is the value of that piece of work? To me, control over means that we can get closer to sensory apparatus, to how people think, feel, communicate, hallucinate and dream.

When the process is somehow unique and rare--meaning either expensive, hard to get access to, or difficult to execute, judgements about the material can easily become clouded. Exploration with new tools is halting at times because users need to form new standards of judgements, and it useful perhaps to remember that in the course of innovation, the world may end up looking differently than it did before. In my case, I found that holding on to some aspects of the familiar in the picture domain--the unprocessed images, a montage style which bowed a little to narrative conventions--allowed me to work with less concrete images in the soundtrack.

> *A miracle is always a transformation of Nature.*
>
> -Barthes

ILLUSTRATIONS

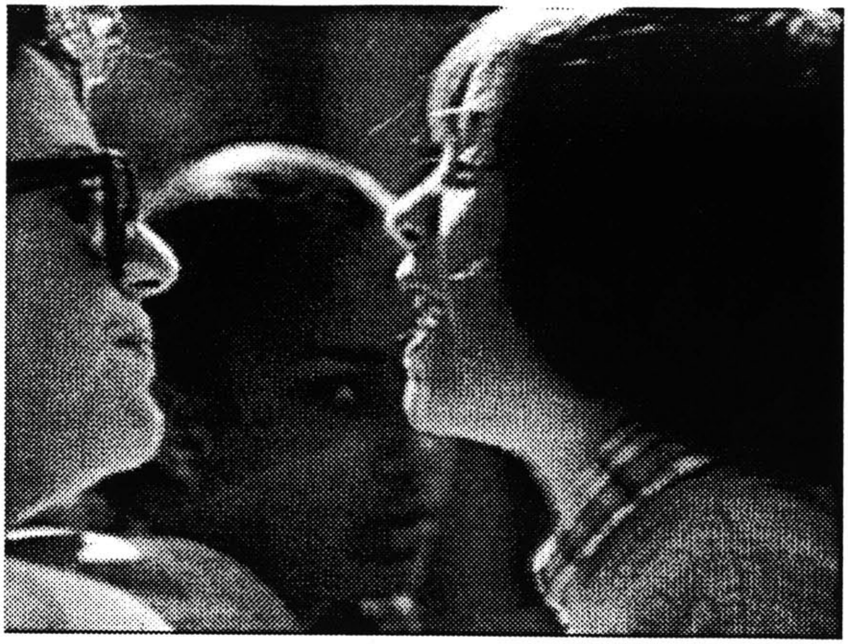


1.)

2.)

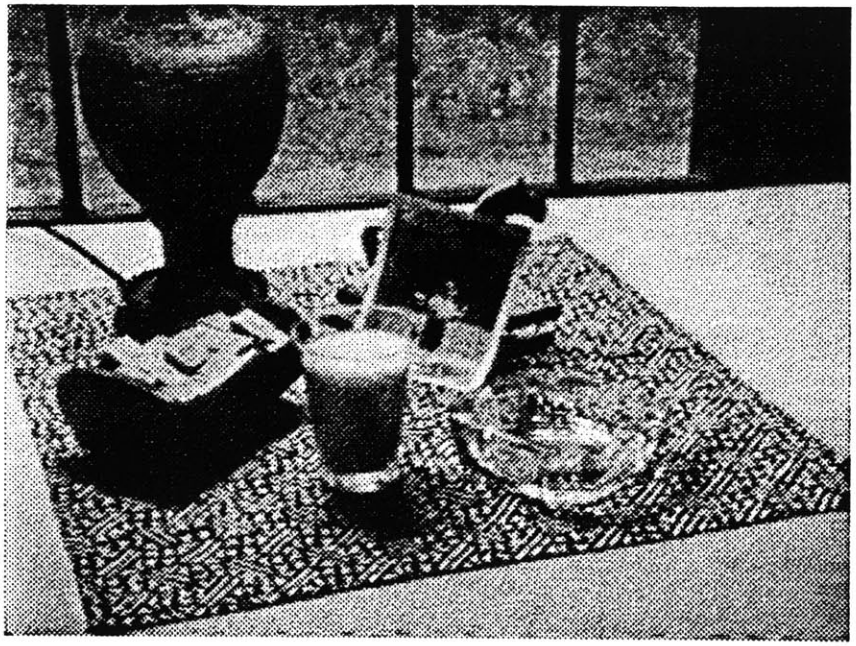
3.)

4.)

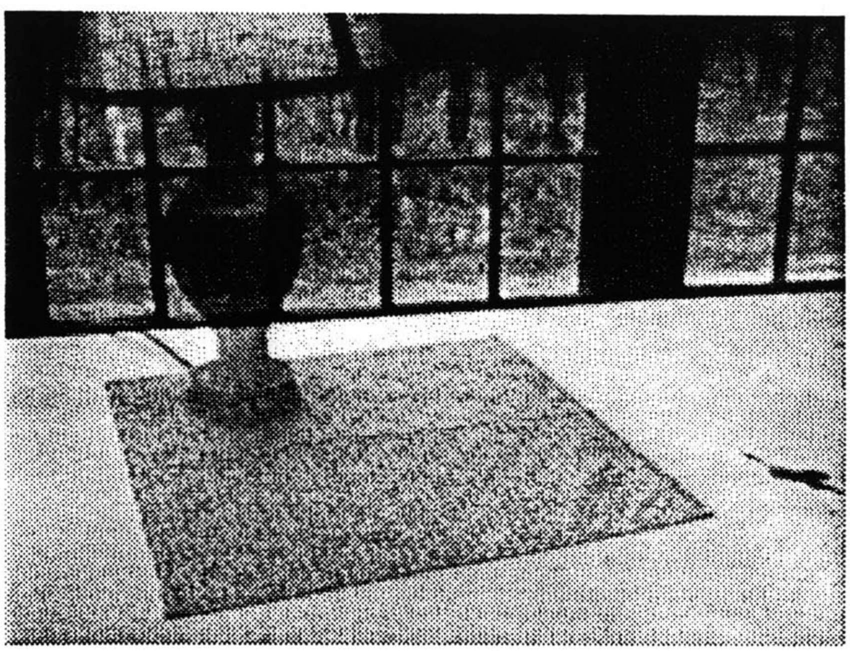


5.)

6.)

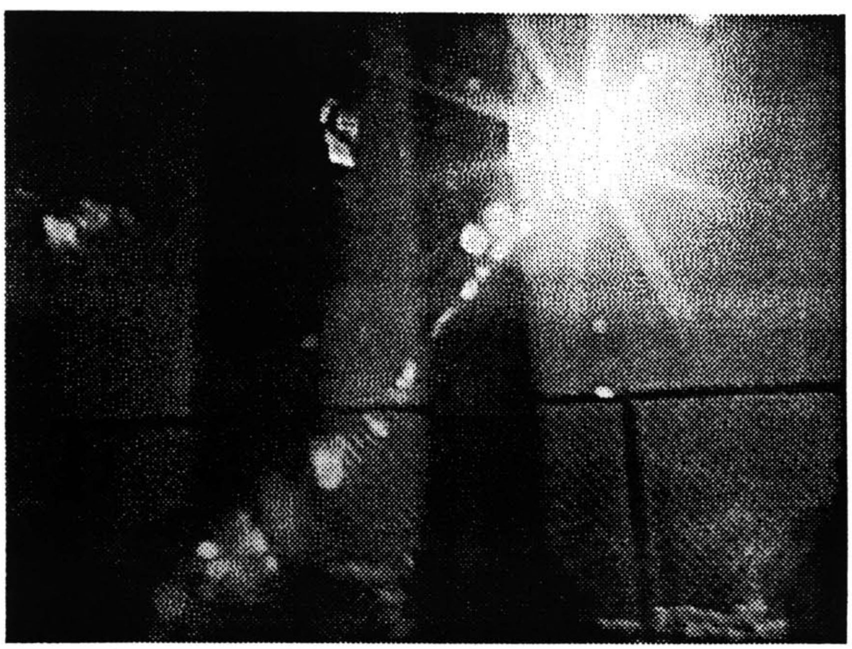


7.)

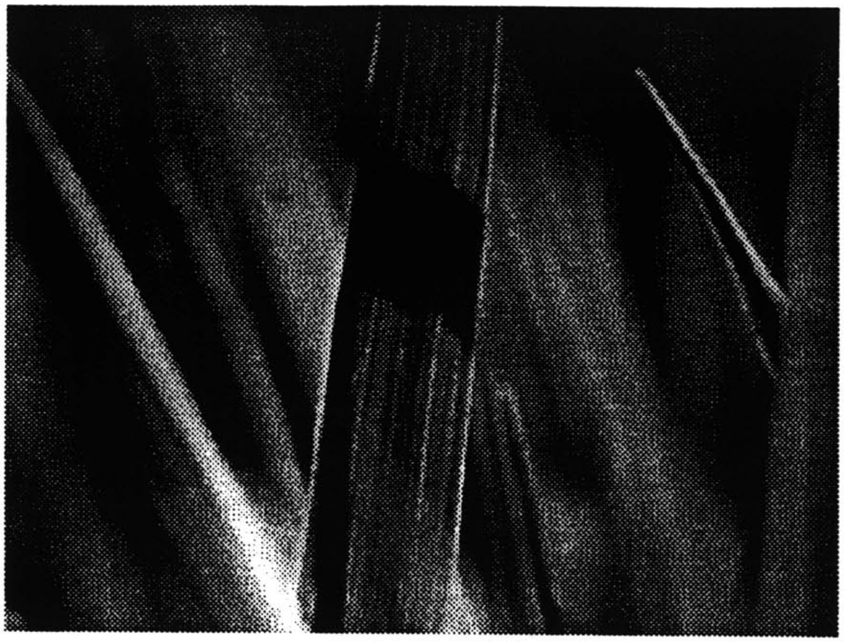


8.)

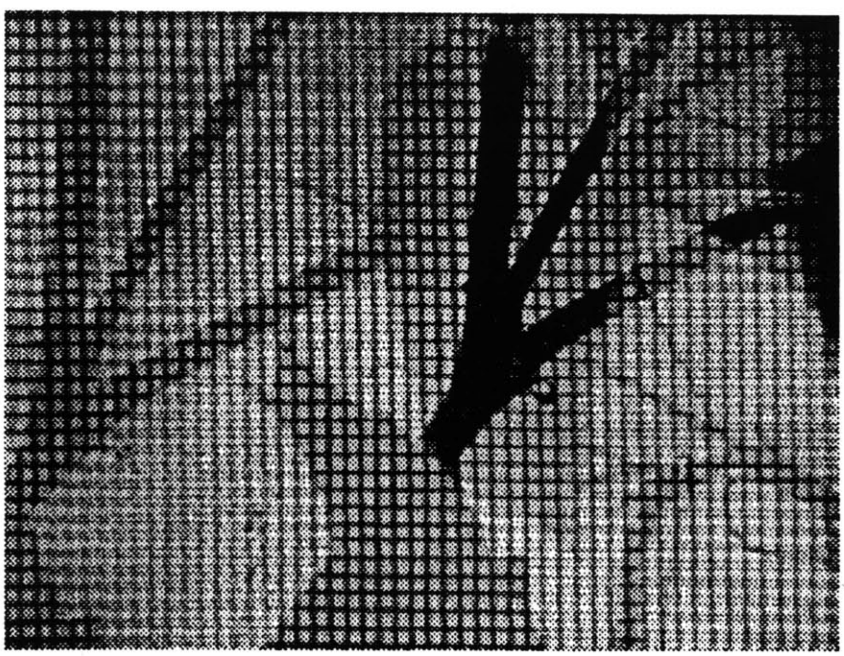
9.)



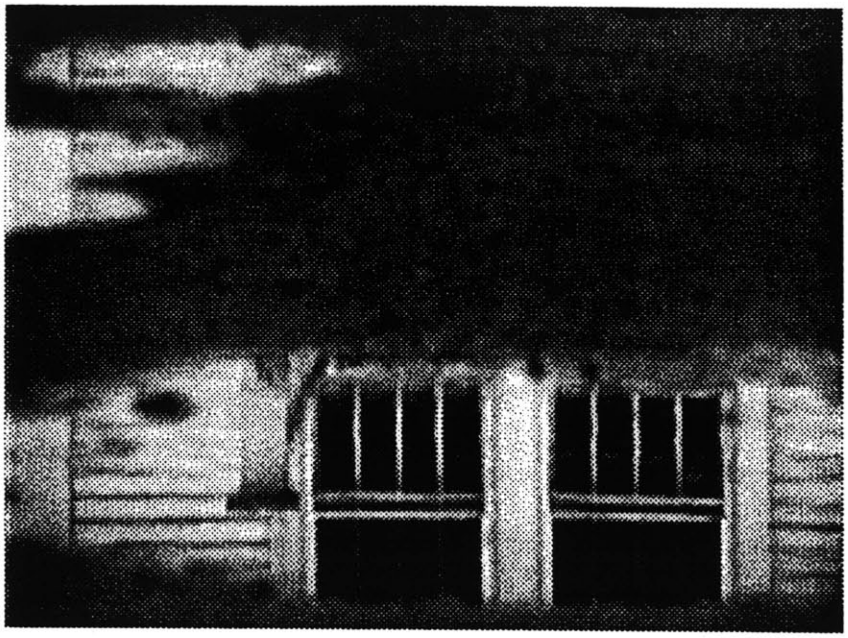
10.)



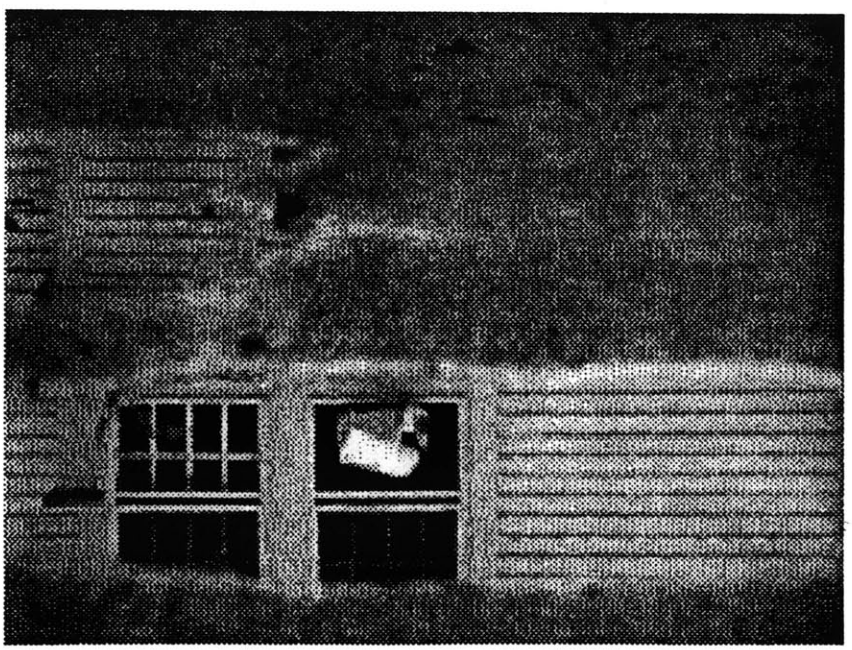
11.)



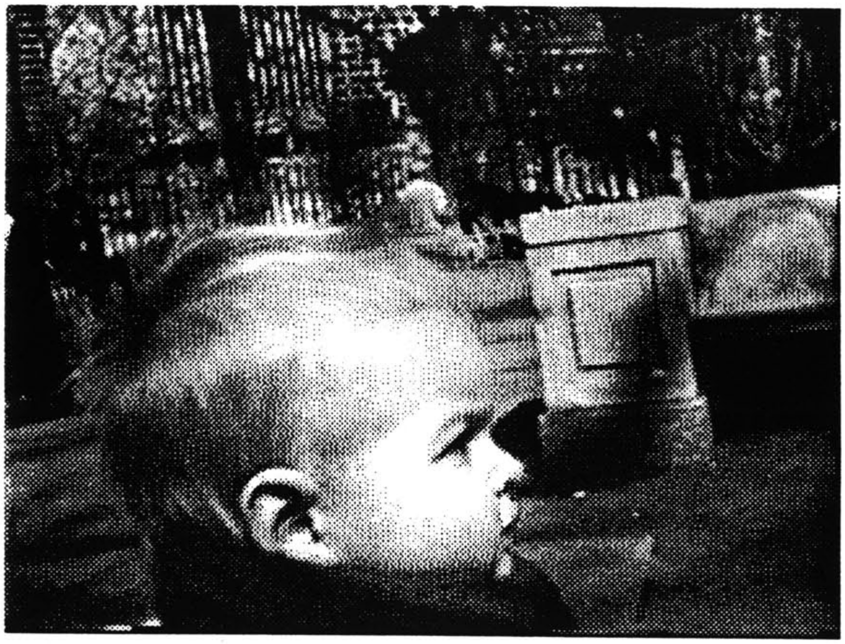
12.)



13.)



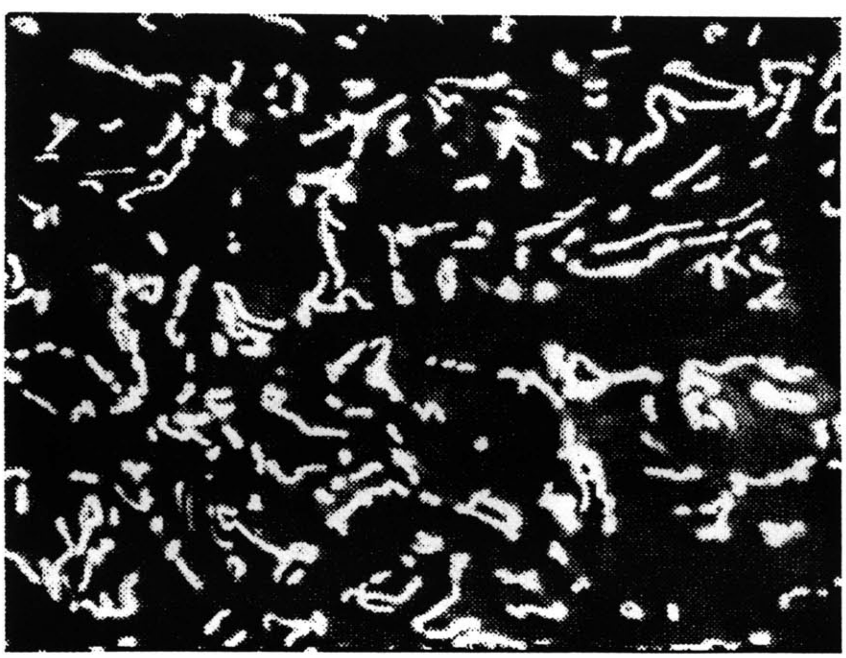
14.)



15.)

16.)

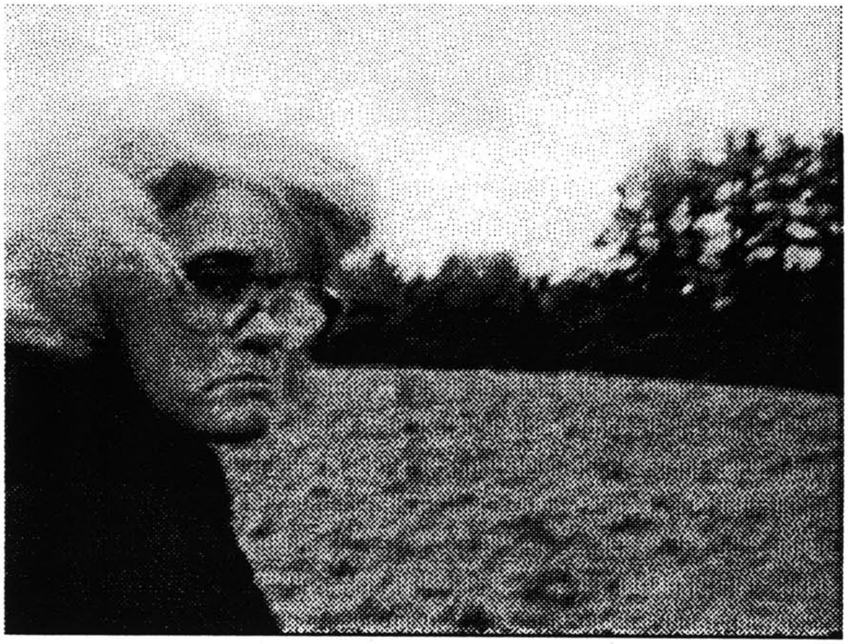
17.)



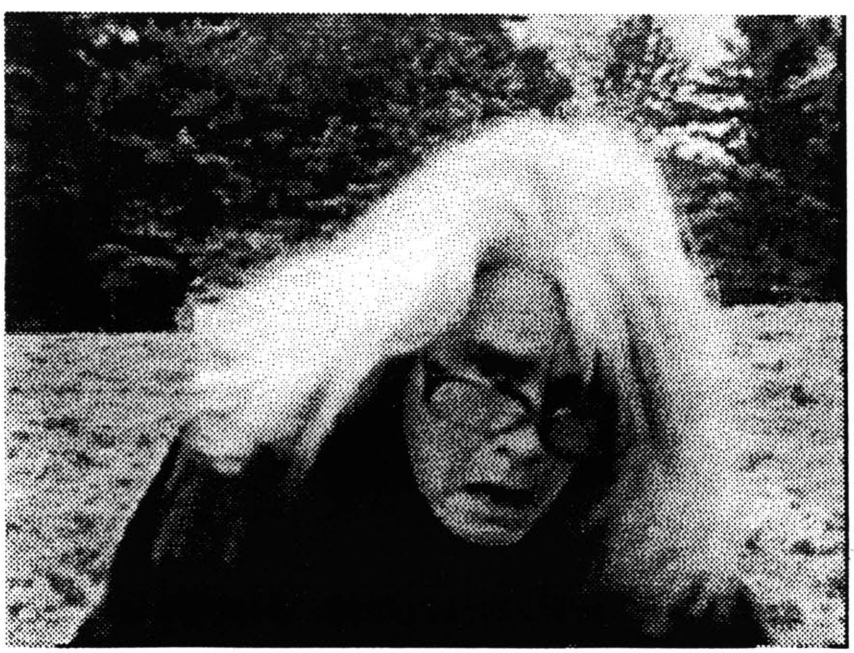
18.)

19.)

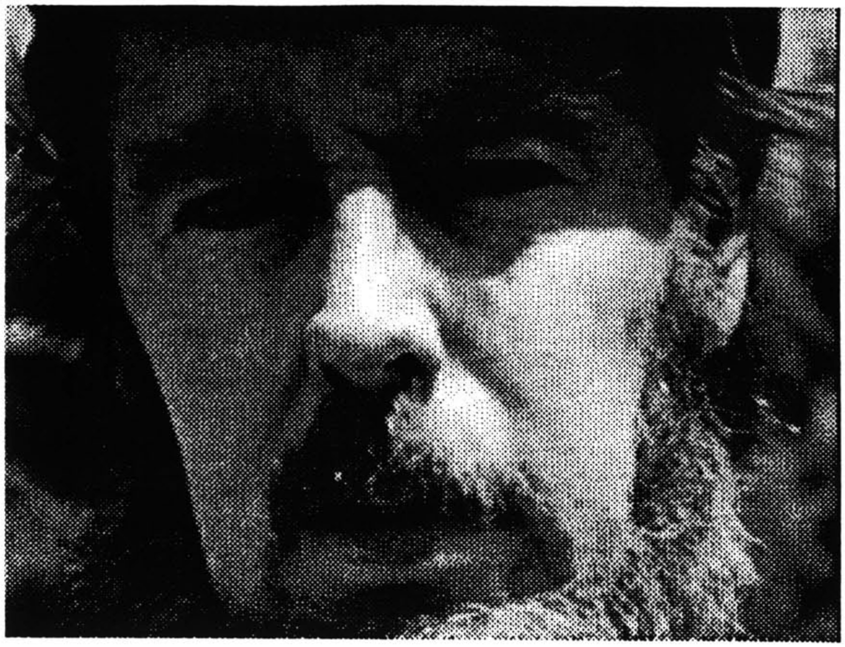
20.)



21.)

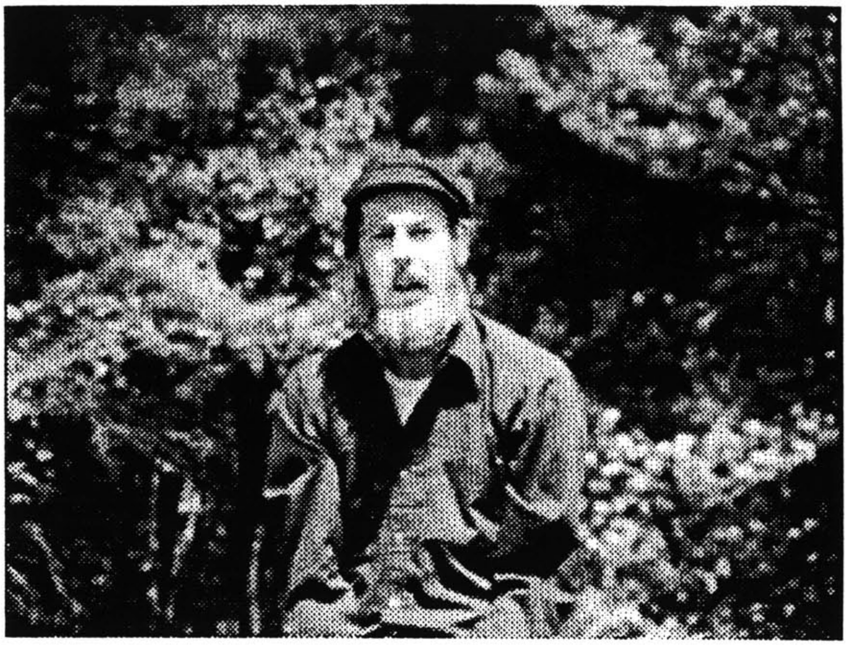


22.)

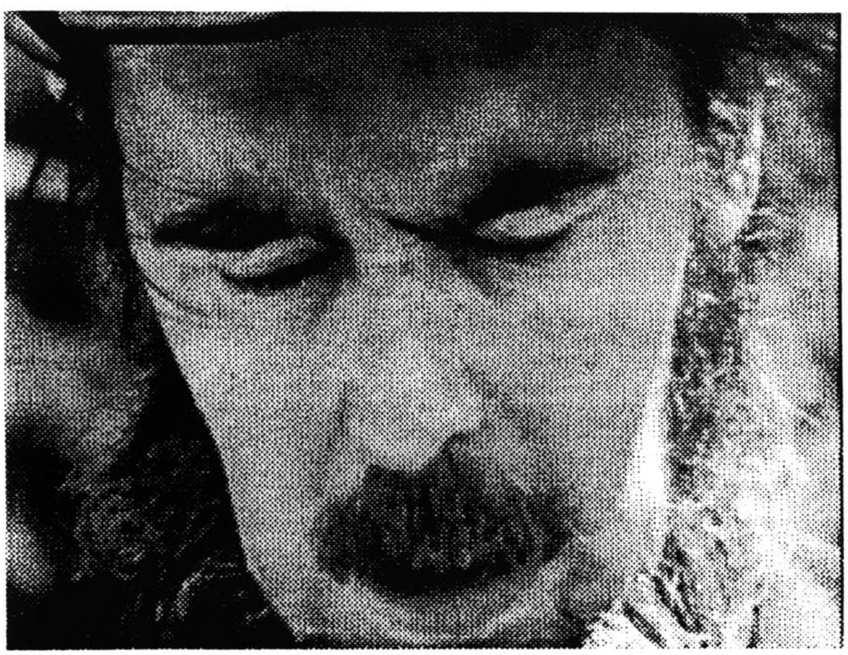


23.)

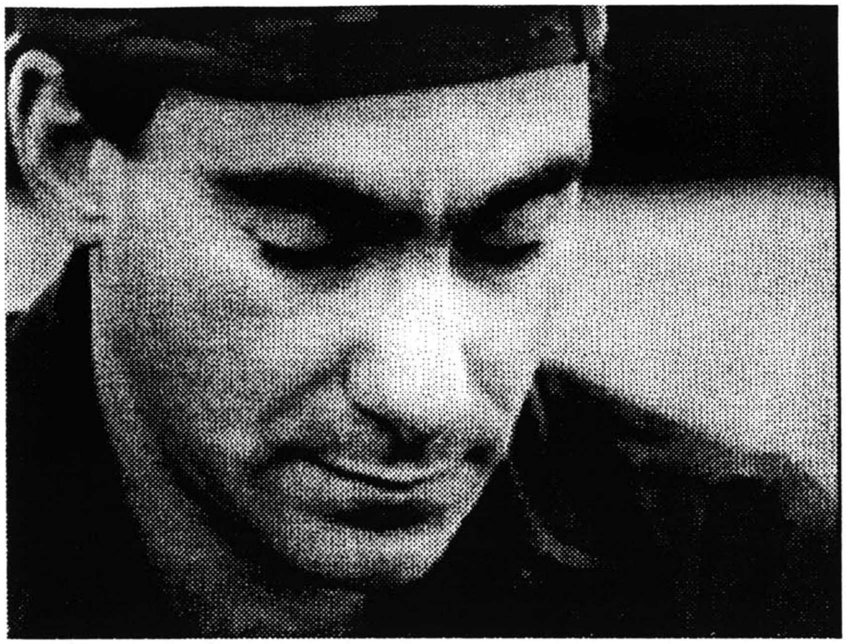
24.)



25.)



26.)



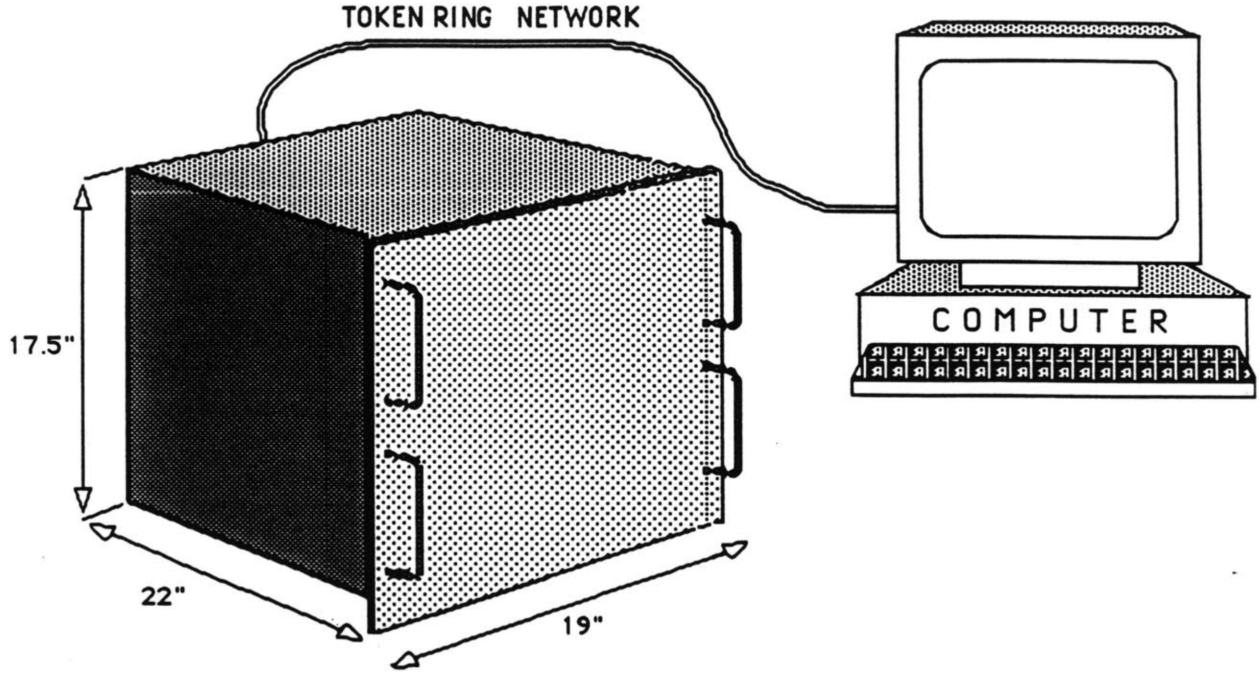
27.)

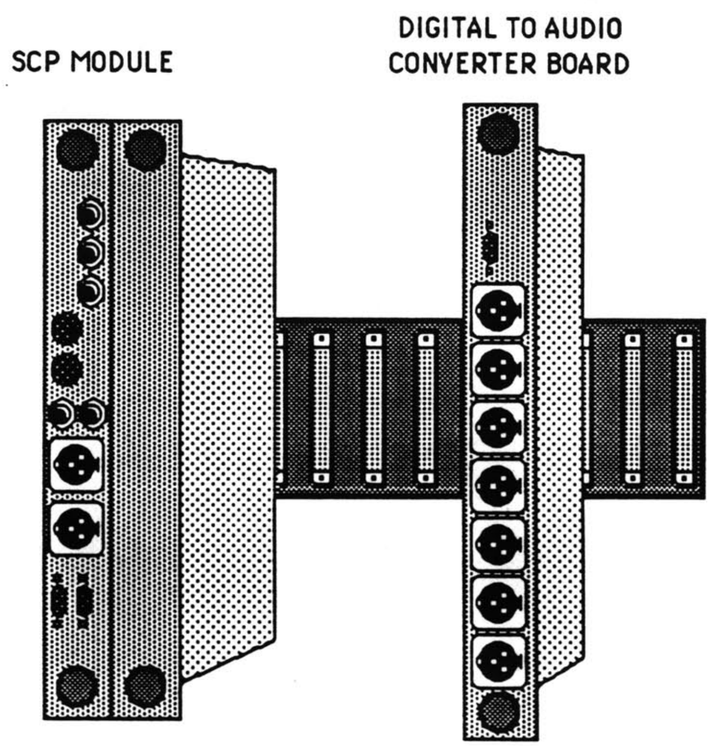
28.)

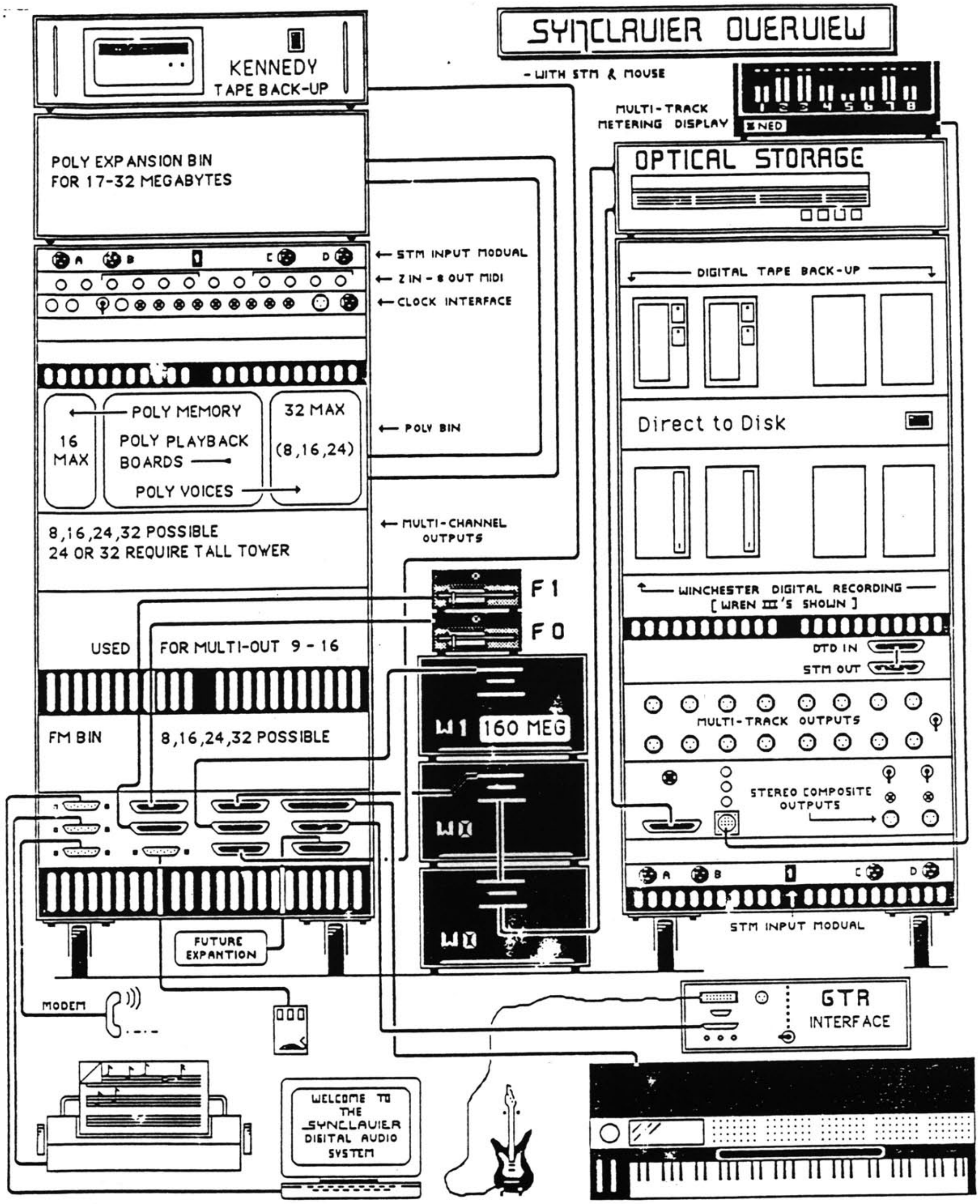
29.)

30.)

THE AUDIOFRAME

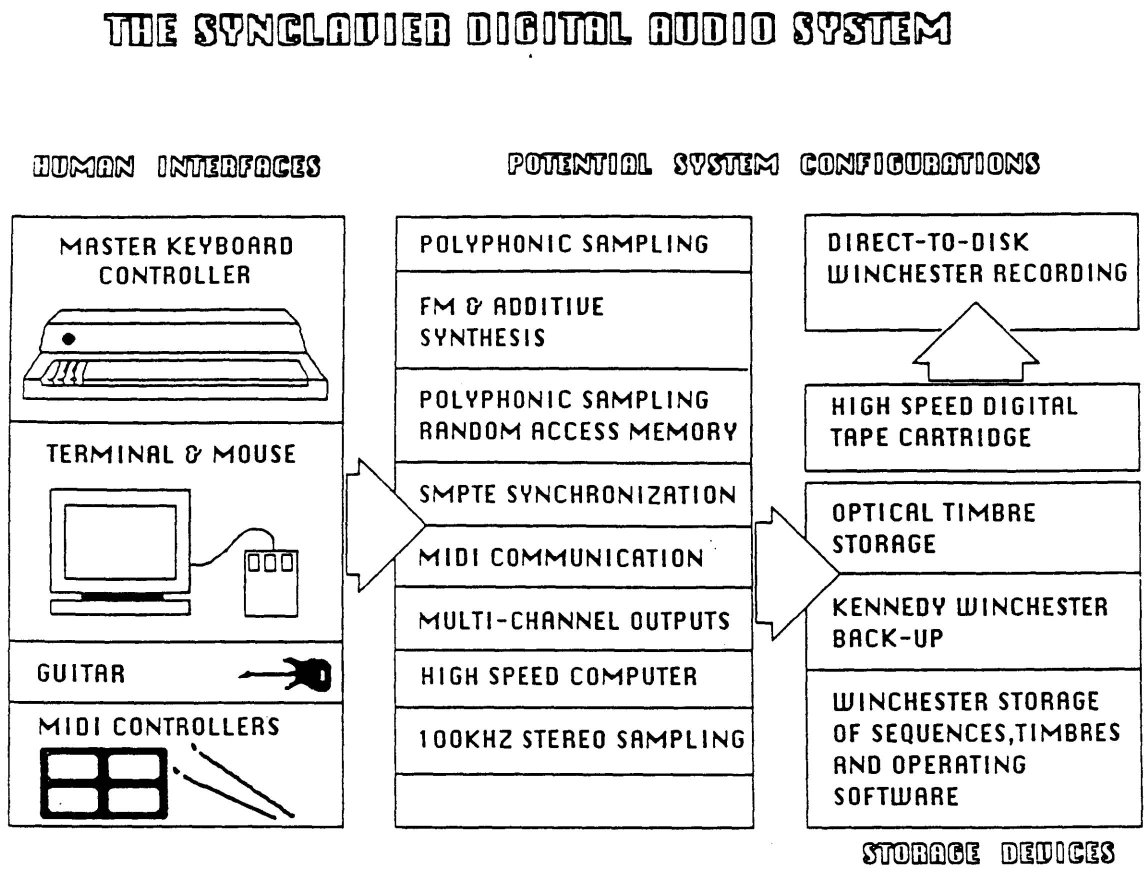





SYNCLAVIER OVERVIEW
- WITH STM & MOUSE
KENNEDY
TAPE BACK-UP
POLY EXPANSION BIN
FOR 17-32 MEGABYTES
STM INPUT MODUAL
2 IN - 8 OUT MIDI
CLOCK INTERFACE
POLY MEMORY
16 MAX
POLY PLAYBACK BOARDS
32 MAX
(8,16,24)
POLY VOICES
POLY BIN
8,16,24,32 POSSIBLE
24 OR 32 REQUIRE TALL TOWER
MULTI-CHANNEL OUTPUTS
USED
FOR MULTI-OUT 9 - 16
FM BIN
8,16,24,32 POSSIBLE
F1
F0
W1 160 MEG
W0
W0
FUTURE EXPANTION
MODEM
MULTI-TRACK METERING DISPLAY
OPTICAL STORAGE
DIGITAL TAPE BACK-UP
Direct to Disk
WINCHESTER DIGITAL RECORDING
[WREN III'S SHOWN]
DTD IN
STM OUT
MULTI-TRACK OUTPUTS
STEREO COMPOSITE OUTPUTS
STM INPUT MODUAL
GTR
INTERFACE
WELCOME TO THE SYNCLAVIER DIGITAL AUDIO SYSTEM

MEMORY STORAGE IN SYNCLAVIER DIGITAL AUDIO WORKSTATION

DESCRIPTION	PURPOSE	SIZE
FLOPPY DISKETTES	-INITIALIZE SYSTEM -SOFTWARE RELEASE UPGRADES	-1.2 MEGABYTES PER FLOPPY
WINCHESTER HARD DISK	-LARGE,HIGHSPEED STORAGE OF TIMBRES & SEQUENCES -DIRECT-TO-DISK RECORDING	-160 & 300 MEGABYTES PER DRIVE
HIGH SPEED DIGITAL TAPE BACK-UP CARTRIDGE	-BACK-UP FOR DIRECT-TO-DISK RECORDING	-240 MEGABYTES PER CARTRIDGE
KENNEDY TAPE TIMBRE STORAGE	-BACK-UP FOR WINCHESTER DRIVES	-15 MEGABYTES PER CARTRIDGE
HIGH CAPACITY OPTICAL STORAGE	-VERY HIGH CAPACITY ONLINE STORAGE FOR SYNCLAVIER TIMBRES	-2000 MEGABYTES PER DISK
POLYPHONIC SAMPLING RANDOM ACCESS MEMORY	-REALTIME ACCESS TO SYNCLAVIER TIMBRES	-4 TO 32 MEGABYTES
SEQUENCER RANDOM ACCESS MEMORY	-REALTIME STORAGE AND ACCESS TO SYNCLAVIER NOTELIST AND ALL SEQUENCER EVENTS	-512K PER CARD -UP TO 8 MEG'S


THE SYNCLAVIER DIGITAL AUDIO SYSTEM
HUMAN INTERFACES
POTENTIAL SYSTEM CONFIGURATIONS
MASTER KEYBOARD CONTROLLER
TERMINAL & MOUSE
GUITAR
MIDI CONTROLLERS
POLYPHONIC SAMPLING
FM & ADDITIVE SYNTHESIS
POLYPHONIC SAMPLING RANDOM ACCESS MEMORY
SMPTE SYNCHRONIZATION
MIDI COMMUNICATION
MULTI-CHANNEL OUTPUTS
HIGH SPEED COMPUTER
100KHZ STEREO SAMPLING
DIRECT-TO-DISK WINCHESTER RECORDING
HIGH SPEED DIGITAL TAPE CARTRIDGE
OPTICAL TIMBRE STORAGE
KENNEDY WINCHESTER BACK-UP
WINCHESTER STORAGE OF SEQUENCES,TIMBRES AND OPERATING SOFTWARE
STORAGE DEVICES

BIBLIOGRAPHY

Abbado, Adriano. "Perceptual Correspondences of Abstract Animation and Synthetic Sound." M. S. thesis. Cambridge, Massachusetts: Massachusetts Institute of Technology, 1988.

Anderton, Craig. MIDI for Musicians. New York: Amsco Publications, 1986.

-----. "The AudioFrame Grows Up." MIX. January 1989, 126-131.

Barthes, Roland Camera Lucida. Translated by Richard Howard. New York: Hill and Wang, 1981.

Barthes, Roland Mythologies. Selected and translated by Anette Lavers. New York: Hill and Wang, 1972.

Berger, John. Ways of Seeing.

Bettelheim, Bruno. The Uses of Enchantment. New York: Vintage Books,1977

Bjorkman, Stig; Manns, Torsten; and Sima, Jonas, eds. Bergman on Bergman. Translated by Paul Britten Austin. New York: Simon and Schuster, Inc., 1973.

Borges, Jorge Luis. Seven Nights. Translated by Eliot Weinberger. New York: New Directions Publishing Corporation, 1984.

Borges, Jorge Luis. A Personal Anthology. Edited by Anthony Kerrigan. New York: Grove Press, Inc., 1967.

Bunuel, Luis. My Last Sigh. Translated by Abigail Israel. New York: Vintage Books, 1984.

Cameron, Evan William, ed. Sound and the Cinema: The Coming of Sound to American Film. Pleasantville: Readgrave Publishing Co., 1980.

Eisenstein, Sergei. Film Form. Edited and translated by Jay Leyda. New York: Harcourt Brace Jovanovich, Inc., 1949.

Eisenstein, Sergei. The Film Sense. Edited and translated by Jay Leyda. New York: Harcourt Brace Jovanovich, Inc., 1975.

Emmerson, Simon, ed. The Language of Electroacoustic Music. New York: Harwood Academic Publishers, 1986.

Fox, Jordan. "Walter Murch--Making Beaches Out of Grains of Sand." Cinefex. 3 (December 1980), 42-57.

Frampton, Hollis. Circles of Confusion. Rochester: Visual Studies Workshop Press, 1983.

Frederick, David and Pentell, Laurie. Audio Frame Reference Manual. Colorado: Wave Frame Corporation, 1988.

Gentry, Rick. "Alan Splet and Sound Effects for *Dune*." American Cinematographer. 65:11 (December 1984), 62-72.

Godard, Jean Luc. Godard on Godard. Translated and edited by Tom Milne. New York: Da Capo Press, 1972.

Griffiths, Paul. Modern Music, The Avant-Garde since 1945. New York: George Braziller, 1981.

Harvey, Jonathan. *Mortuos Plango, Vivos Voco*: A Realization at IRCAM." Computer Music Journal. 5:4(Winter 1981), 22-24.

Kuhn, Thomas S. The Structure of Scientific Revolutions. 2nd ed., enlarged. Chicago: University of Chicago Press, 1979.

Machover, Tod. "A View of Music at IRCAM." Contemporary Music Review. 1:1 (1984), 1-10.

Mast, Gerald and Cohen, Marshall, eds. Film Theory and Criticism. 2nd ed. New York: Oxford University Press, 1979.

McAdams, S. "Music: A Science of the Mind?". in Contemporary Music Review, S. MacAdams ed. London: Gordon and Breach, 1987

McAdams, S. and Saariaho, K. "Qualities and Functions of Musical Timbre". in Proceedings, International Computer Music Conference, 1885.

Meyer, Chris. "The Audioframe Explained." Part two: the present. Music and Technology. October 1988, 44-47.

-----. "The Audioframe Explained." Part three: mixing and recording. Music and Technology. November 1988, 34-39.

-----. "The Audioframe Explained." Part four: sound and modeling. Music and Technology. December 1988, 48-52.

Michelson, Annette, ed. Kino-Eye, the Writings of Dziga Vertov. Translated by Kevin O'Brien. Berkeley: University of California Press, 1984.

Miller, Alice. The Drama of the Gifted Child. Translated by Ruth Ward. New York: Basic Books, Inc., 1981.

Minsky, Marvin. The Society of Mind. New York: Simon and Schuster, 1986.

-----. "Music, Mind, and Meaning." Computer Music Journal. 5:3 (Fall 1981), 28-44.

Monaco, James. How to Read a Film. New York: Oxford University Press, 1977.

Moorer, James A. "The Lucasfilm Audio Signal Processor." Computer Music Journal. 6:3 (Fall 1982), 22-32.

Moorer, James A. *et al.* "The 4C Machine." Computer Music Journal. 3:3, 16-24.

Neale, Steve. Cinema and Technology: Image, Sound, and Colour. Bloomington: Indiana University Press, 1985.

Newhall, Beaumont. The History of Photography. New York: The Museum of Modern Art. 1982.

Norman, Donald. The Psychology of Everyday Things. New York: Basic Books, Inc., 1988.

Olesha, Yuri. Envy and Other Works. Translated, with an introduction, by Andrew R. MacAndrew. New York: W.W. Norton and Co., 1967.

Pine, Ted. "The Tapeless Studio (R): Its Concept, Development, and Applications." An in-house systems history. New England Digital Corporation.

"Questions and Answers About SoundDroid." Promotional Material for *The Droid Works*, an affiliate of Lucasfilm, Ltd. and Convergence Corporation. 4/9/86.

Riesz, Karel and Millar, Gavin. The Technique of Film Editing. 2nd ed. London: Focal Press, 1968.

Risset, Jean-Claude. "Computer Music Experiments 1964-..." Computer Music Journal. 9:1 (Spring 1985), 11-18.

Roads, Curtis, ed. Composers and the Computer. Los Altos, California: William Kauffman, Inc., 1985.

-----. The Music Machine. Cambridge, Massachusetts: The MIT Press, 1979.

-----. "A Conversation with James A. Moorer." Computer Music Journal. 6:4 (Winter 1982), 10-21.

Rockwell, John. All American Music: Composition in the Late Twentieth Century. New York: Vintage Books, 1984.

Russett, Robert and Starr, Cecile. Experimental Animation. New York: Van Norstrand Reinhold Company, 1976.

Sartre, Jean-Paul. Being and Nothingness. Special abridged edition. Translated by Hazel E. Barnes. Secaucus: The Citadel Press, 1956.

Sitney, P. Adams. Visionary Film: The American Avant-Garde, 1943-1978. 2nd ed. New York: Oxford University Press, 1979.

Snell, John. "The Lucasfilm Real-Time Console for Recording Studios and Performance of Computer Music." Computer Music Journal. 6:3 (Fall 1982), 33-45.

Sontag, Susan. On Photography. New York: Dell Publishing Co., Inc., 1977.

Synclavier. Operations manuals. New England Digital Corporation, 1988.

Tarkovsky, Andrei. Sculpting in Time. Translated by Kitty Hunter-Blair. New York: Alfred K. Knopf, Inc., 1987.

Weis, Elisabeth and Belton, John. Film Sound. New York: Columbia University Press, 1985.

Wishart, Trevor. On Sonic Art. York: Imagineering Press, 1985.