SIMULATING THE CUES OF SPATIAL HEARING IN NATURAL ENVIRONMENTS

Gary S. Kendall and William L. Martens
Computer Music Studio
Northwestern University
Evanston, IL 60201

ABSTRACT

In sound reproduction the apparent direction from which a sound emanates is typically controlled by shifting the balance of sound between the speakers. This technique often produces diffuse images and is only capable of locating sound sources between the speakers. In the last decade we have come to recognize that in addition to interaural intensity and time differences there is a localization cue provided by the reflection of sound off the convolutions of the pinna, the shoulders and the upper torso. Measured head-related transfer functions vary considerably from one individual to the next, but there are common trends in how spectral features change with the sound source angle of incidence. These directional transfer functions can be modeled, simplified, and enhanced by a digital filtering network so as to produce idealized spectral cues for localization. Imposing these idealized spectral cues on arbitrary signals such as those in music can fool the auditory system and induce three-dimensional spatial images even when reproduced with stereo loudspeakers.

Our experience with simulated spatial cues has led us to the initial design of a reverberator with different goals than those of Schroeder (1962) or Moorer (1979). The primary goal of our design is the production of "spatial reverberation," a spatially distributed reflected sound field intended to help listeners localize sounds while imparting a very strong impression of a specific reverberant environment. For this reason, our "spatial reverberator" can be configured so as to replicate nearly exactly the spatial and temporal distribution of reflected sound for a modeled room. The design includes multiple outputs which can be separately directionalized with our simulated pinna transfer functions to produce spatially positioned reverberation "streams." The "spatial reverberator" captures both spatial and temporal changes in the distribution of reflected sound that occur with changes in the position of the sound source, and even with changes in the position of the listener in the modeled room. Also considered in our design are the subjective attributes of reverberation described in the literature of subjective room acoustics, such as "definition" and "spatiousness" (Rasch and Plomp, 1982). Since the spatial and temporal distribution of reflected sound determines these subjective attributes as well as distance, the "spatial reverberator" can provide direct control over a wide range of spatial percepts important for sound processing and reproduction.

I. INTRODUCTION

When a sound event is transduced into electrical energy by a microphone and reproduced over loudspeakers or headphones, the experience of the sound event is altered dramatically from what would result if the listener were located at the position of the microphone. One of the primary reasons for the change in the experience is that information regarding the spatial location of the sound event and of the sound reflected from the environment has been lost. Multi-channel recording and reproduction can retain some spatial information, but conventional techniques do not attempt to recreate the spatial sound field of a natural environment and, therefore, create a listening experience which is spatially impoverished. This is despite the fact that most recordings include reverberation from a natural environment in order to provide the listener with a general impression of an acoustic environment. There is a need for improved techniques to spatialize recorded sound, but the need is even more urgent for synthesized sounds which have no spatial attributes save those provided by signal processing.

The original goal of the work described in this paper was to provide computer music composers with a comprehensive control of auditory space when compositions are created for headphone listening. It was assumed that the cues used by the auditory system in forming spatial images could be replicated accurately only with headphones. We quite accidentally found this not to be the case. Quite often, reproduction with speakers was just as vivid as headphones. It was also assumed that our original goal could be achieved by simulating cues for directional hearing and then adding reverberation. This also proved not to be the case. In fact, we could not have forseen the complexity of the issues involved and the degree to which the many different components of spatial hearing in natural environments interact in the formation of spatial images.

The literature of spatial hearing in all of its facets is distributed in many different journals. Most researchers pursue only one narrow segment of the literature. Psychoacoustic research has concentrated on three categories of cues for directional hearing: interaural intensity differences, interaural time differences, and spectral cues introduced by the pinna, head, and torso. Research into the first two types of cues had produced most of its major findings by the 1960's while research into spectral cues is today still a quite active area. All of this research has treated these topics as separate issues and assumed that the auditory system's mechanisms could be studied in isolation. This assumption is valid insofar as one intends to study "separate mechanisms" but it is clear that in natural listening situations all mechanisms work together in the formation of spatial images.

"Spatial hearing" in natural environments is a broader topic than this concentration of psychoacoustic research itself would indicate. What has commonly been referred to as "localization" research actually is concerned only with directional hearing. The perceived location of a sound source in a natural environment has two relatively independent dimensions - direction and distance. Distance perception has received substantially less attention than directional hearing for two reasons. First, the "mechanism" for distance perception appears to be less clearly located in or near the peripheral auditory system and the modeling of "higher-level" auditory mechanisms is harder to conceptualize. Second, it is extremely difficult to perform experiments that control a number of spatial hearing cues simultaneously, especially if those cues are going to be realistic. Distance is largely perceived on the basis of reflected sound from the environment. But reflected sound also constitutes the basis for several other perceptual qualities involved with spatial hearing. These are perceptual qualities of space investigated in a discipline which has come to be called "subjective room acoustics." "Definition" and "spatiousness" are two of the terms often employed to describe the subjective attributes of how sound images are perceived in the context of rooms. We question whether it is possible to study localization in the absence of these subjective properties. Spatial hearing experiments conducted without reflected sound only study the responses of the auditory system under conditions of impoverished information.

Since our goal is to provide composers with a comprehensive control of auditory space percepts in music, we must approach the problem of simulating spatial cues as a whole. Our effort in this regard has had two primary components. The first is the formulation of idealized spectral cues for use in directionalizing sound. We know, for example, that spectral cues induce spatial percepts even when other types of cues are absent. The second is the simulation of environmental reverberation that retains the spatiality of reflected sound. By combining spectral cues for directional hearing with such reverberation, we are attempting to recreate the experience of listening in natural environments entirely from computer simulation. We use the term "spatial reverberation" for this synthesis of directional cues and simulated reflected sound. It is our hope that techniques like ours will stimulate composers to produce a kind of music that not only takes place in space but is spatially conceived.

II. DIRECTIONAL HEARING

A. Interaural Differences

The two cues which classical psychoacoustics holds as primarily responsible for identifying the direction or incidence angle of a sound source are interaural intensity difference (IID) and interaural time delay (ITD). The physical basis for these cues is as follows: As a sound source moves on the horizontal plane toward the side - away from directly ahead or directly behind the listener - IID grows from 0 dB up to roughly 20 dB depending on frequency, and the ITD grows from 0 to about 650 microseconds (Feddersen, et al., 1957). Because the head blocks only those frequencies with wavelengths shorter than the diameter of the head (about 1 KHz and above), the acoustic "head shadow" responsible for the IID is frequency dependent. For pure-tones, IID is only a salient cue at frequencies higher than 1 KHz. Because the periods of high frequencies are shorter than the maximum ITD, ITD is a salient cue for pure-tones only below about 1.5 KHz. The potential confusion between waveform periods and ITD's is abolished if the high frequency stimuli have time-varying amplitude (eg., Nuetzel and Hafter, 1976) or time-varying frequency (eg., Blauert, 1981).

In actuality, IID and ITD only provide the auditory system with information on whether a sound source is to the left or right of a listener. This is especially clear in headphone listening when cues to externalize the sound source are eliminated (Sakamoto, Gotoh, and Kimura, 1976). Sound images with IID and ITD cues are perceived on a left/right axis inside the head. This is referred to as "lateralization" and has been the subject of considerable research in the last century. The fact that listeners have some basis for identifying the direction of sound sources above, below, in front and in back did not become a general research topic in psychoacoustics until the late 1960's. In order to visualize a point on this left/right lateralization axis projected into three dimensional space, we must imagine a plane containing the point placed perpendicular to the axis. For a given distance from the listener represented by a sphere, the plane cutting through the sphere makes a circle as shown in Figure 1. This circle is a representation of the possible locations in three-dimensional space at which the lateralized sound might have originated.

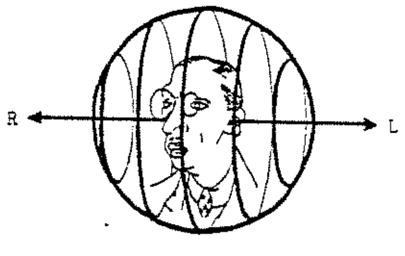

Figure 1 The left/right lateralization dimension and the up/down/front/back circles upon which sounds can be located at a given distance from the listener.

B. Spectral Cues

Within the last 15 years, we have come to recognize that an additional cue for directional hearing is provided by the reflection of sound off the convolutions of the pinna (outer ear), shoulders and the upper torso. These short latency reflections impose directional information on the spectrum of the source signal. The most important of these reflections are those contributed by the pinnae. Because the pinnae have a very asymmetric arrangement of ridges, the composite sound reflections create a unique spectral profile for every sound direction. The auditory system uses these spectral profiles to remove the spatial ambiguity that results from IID and ITD cues alone. It is the spectral cues produced by the pinna, shoulders and upper torso that enable the auditory system to determine the position of the sound source on the circle represented in Figure 1– above, behind, below or in front of the listener.

Numerous researchers have studied the relationship between the direction of a sound source and the acoustic transformation produced by the pinna. It is a reasonable approximation to imagine that a discrete reflection from a ridge of the pinna will produce a notch in the spectrum of the source signal and that a collection of reflections will produce a complex spectral shape with many notches of varying depth. Even though some empirical measurements of the pinna made in the ear canal conform to this approximation, most measurements indicate that reflections are not discrete and that there is no simple correspondence between single reflections and spectral notches. Understanding of the relationship between time-domain and frequency-domain representations of the pinna responses was historically slow to evolve. Batteau (1967) was the first researcher to develop a model of the relationship between the physical characteristics of the pinna and the acoustic information used for directional hearing. He attempted to identify individual reflections responsible for judgements of azimuth and for elevation. Shaw and Teranishi (1968) measured the acoustic effects of the pinna in the frequency domain and demonstrated that the relationship between pinna filtering and the directional location of a sound source is very complex, but they did not attempt to explain how the auditory system used this information. Blauert (1969) hypothesized that pinna cues were evaluated by the auditory system in terms of spectral "preference bands" that constituted a signature of certain source directions. It was not until 1974 that Wright, Hebrank and Wilson synthesized Batteau's and Blauert's views into a concept that bridges the time-domain/frequency-domain distinction and demonstrates the essential similarity of both views.

Despite the fact that pinna transfer functions are highly complex and difficult to relate to a simple model, they have easily identified spectral features. A quick examination reveals that they contain spectral notches and peaks whose frequencies are dependent on the incidence angle of the source signal. Almost every researcher has noted that individual pinna transfer functions vary tremendously from each individual ear to the next. Careful examination reveals that despite the variety of details there are numerous common trends. For example, on the lateral plane (the plane defined by the left/right dimension and the above/below dimension), the frequencies of the two most prominent spectral notches generally increase with increasing elevation. The exact shapes of the head-related transfer functions can differ as shown here for subject MDL (Figure 2a) and subject GSK (Figure 2b), but both show the same trend in the migration of these spectral notches. This might suggest that the directional information supplied by the pinna can be largely characterized in terms of these spectral notches, although there are several other observable global trends involving spectral peaks and overall spectral contour. In fact, one can separate to some extent the individual spectral features contributed by the head and the pinna. Binaural recordings and studies such as that by Butler and Belendiuk (1977) demonstrate that it is quite possible for one person to utilize the spatial hearing cues recorded with another person's ears, but the issue of how the auditory system evaluates the complex spectral profile at the two ears has not been adequately investigated and may require many more years of research.

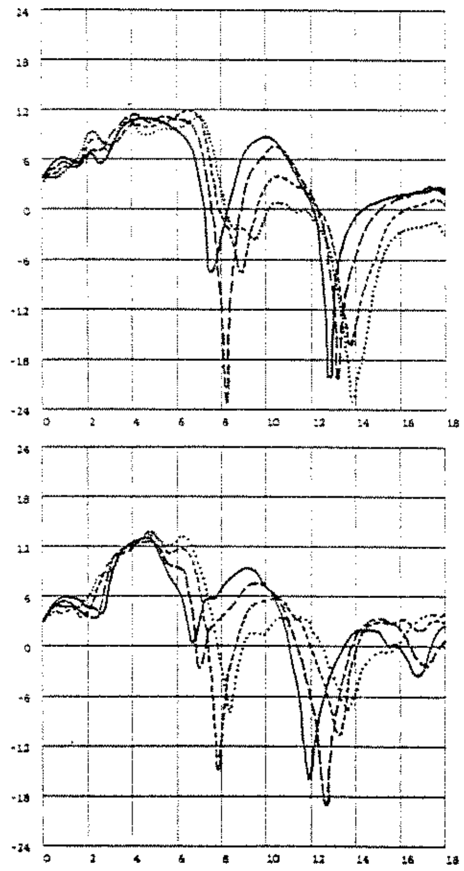

Figure 2 Head-related transfer functions for two subjects illustrating trends in spectral features for increasing source elevation in the lateral plane. The sound was located 2 meters to the left of the subject and was moved from ear level (0 degrees) to an elevation 30 degrees above ear level (solid line – 0 deg., long dashes – 10 deg, short dashes – 20 deg., and dotted line – 30 deg.).

C. Simulating Cues for Directional Hearing

Much of the current research into directional hearing has been conducted with synthetic cues, rather than with actual pinna recordings. Usually, these cues have been synthesized by time-delay networks capturing the essential notches in pinna transfer functions, but ignoring much of the detail which is characteristic of empirical measurements. Stimuli generated in this fashion have been quite useful to research and have often allowed subjects to localize with accuracy near to that of normal listening situations. The most important fact to be learned from this is that the auditory system does not require the complex, "natural" spectral profile in order to correctly identify the direction of a sound source. Simplified spectral cues may be quite sufficient if they include the necessary spectral information (an early attempt at specifying these spectral profiles may be found in Kendall and Rodgers, 1982).

For the purposes of simulating cues for directional hearing in computer music or any kind of audio reproduction, one wishes to determine a set of "idealized transfer functions" that will provide the best possible image of the sound direction for the general population. It has already been established that measured pinna responses display tremendous variability at the detail level, but that in spite of this variability, individuals are able to localize sounds recorded through the pinnae of others. Idealized transfer functions can not be determined by averaging techniques because directional judgements may well be based on spectral features that averaging would smooth away. Notches may appear to be at different frequencies for different individuals, but all individuals have notches. Then too, it may well be possible to create spectral cues that produce superior directional images to those associated with "natural" cues. In this regard, research by Butler and Belendiuk (1977) is supportive. They demonstrated that "some pinnae, . . . , provide more accurate cues for MSP [median sagittal plane] localization than do others." Butler and Belendiuk did not attempt to identify experimentally the exact features which improve directional hearing, but their data did enable them to speculate that "the migration of the notch in the frequency response curves appears more orderly" in the superior pinna.

Our own approach to this problem has been to synthesize directional cues on the basis of considerable visual study of pinna measurements. We have recorded head-related transfer functions for sound sources located 10 degrees apart in azimuth and 20 degrees apart in elevation. We have analyzed our empirical data in order to identify those spectral features which seem the most common to all subjects and most likely to present the auditory system with usable information. Large-scale trends are taken into account in creating a table of spectral features and their relation to perceived direction. It is hoped that by combining the best and most regular features of many different pinna that a set of spectral manipulations can be devised which not only match the saliency of natural cues, but which actually support superior directional perception for many people.

Butler's comment about the "migration of the notch" being "more orderly" was corroborated in the preliminary stages of our own research. When experimenting with artificial transfer functions for three-dimensional space, the first modifications made were to ensure a continuous and orderly progression of spectral changes. Natural pinna functions do not always display this characteristic throughout their range, but it seems to be an important characteristic for synthesized cues. It should be mentioned that Morimoto and Ando (1979) contradict Butler in regard to whether one subject's cues can improve another subject's localization performance, but the three subjects used in their study were chosen to have vastly different pinnae and may well represent the most extreme sort of cases in the general population.

Once we have arrived at a decision on the important spectral features for a given spatial direction, a list of these features is passed on to a filter design program. This program is able to construct a pole-zero filter that matches the prescribed characteristics. The upper panels of Figure 3 show transfer functions measured inside the ear canal for two subjects. The lower panel of Figure 3 shows the transfer function of a digital filter designed to match ideal spectral features derived from such empirical measurements. The problem of designing these filters is somewhat complicated by the fact that one may need to produce continuous changes in direction between the analyzed points. This means that there must be steady migration of poles and zeros as intended direction is changed, or else the resulting signal discontinuities would produce noticeable noise. In effect, the filters must be designed for the entire ensemble of directions taken together.

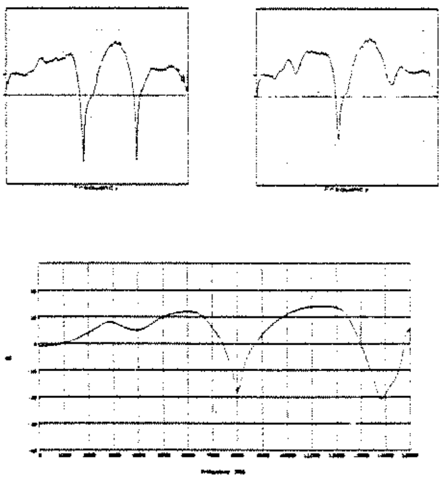

Figure 3 Comparison of measured head-related transfer functions for two subjects (upper panels – maximum is 18 kHz) and our simulation (lower panel – maximum is 15 kHz).

III. SPECTRAL CUES AND LOUDSPEAKER REPRODUCTION

A. The Influence of the Listening Environment

When we first began to experiment with our original pinna simulations in different rooms, we were faced with some rather perplexing experiences. We found that parts of our demonstration tape, especially parts involving the distinction between front and rear images, worked in some situations but not in others. In general, we found that our projected sound paths were deformed even though the general spatial shape was retained. The most revealing demonstration of this phenomenon occurred in a "live end, dead end" or "LEDE" studio monitoring room (Davis and Davis, 1980). A LEDE room is sound absorbent at the end where the speakers are placed and reverberant at the other end behind where the listeners sit. This type of design is preferred today for monitoring rooms in recording studios because it provides excellent stereo imaging for the recording engineer. In this environment our listeners experienced a systematic deformation of the intended sound path depending upon where in the room they listened. As shown in Fig. 4a an intended sound path around the head was easily perceived as such when listening was within the dead-end part of the room. When on the boundary between the live-end and dead-end parts the sound path was deformed but not substantially changed as shown in Fig. 4b. When subjects stepped back into the live-end of the room the sound path ceased to go behind the head but instead made a circular path in front of the listener as shown in Fig. 4c.

Our directional cues were being altered by the acoustics of the listening room. The most obvious factor that changed with our different listening positions was the pattern of reflected sound accompanying the stereo reproduction. Though it is difficult to predict just what acoustic features of a room will result in what sorts of spatial distortion, it is clear that strong early reflections have a significant influence on preceived direction (Barron, 1971). Spectral notches created by these reflections can be mistaken by the auditory system for pinna cues (Rodgers, 1981).

B. Controlling the Listening Environment

Having gained this insight into the source of our spatial distortions, we sought some strategy for bringing the reflected sound of a listening environment under systematic control. To this end we constructed an experimental sound room in which the entire pattern of reflected sound could be controlled through the placement of sound absorbent panels. Our intent was not to create an anechoic environment, but an environment in which reflected sound would be supported in a controlled fashion. We knew that anechoic environments are not only difficult to create but also are terribly oppressive working environments. We wanted to be able to selectively enable or disable discrete reflection paths. We undertook the task of making our room "selectively anechoic", that is, anechoic between the loudspeakers and listening position. Using the technique of "time delay spectrometry" developed by Heyser (1971), we systemati-

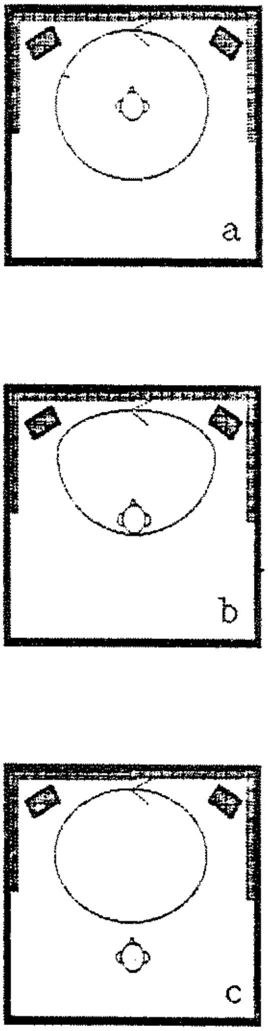

Figure 4 The apparent spatial paths of the sound source on our demonstration tape for three different listening positions in a LEDE environment.

cally eliminated all room reflections that would reach listener from the speakers within the first 20 ms after the arrival of the direct sound. The resulting effect on sound imagery was quite apparent to listeners. Not only were our test tape examples perceived to be following the intended spatial paths, but commercially recorded music was perceived as having greatly improved spatial imagery. (A detailed discussion of the construction and testing of the listening room by Jones, Martens, and Kendall is forthcoming in a separate paper.)

We have discovered that many of our sound examples employing spectral cues are sensitive measures of the quality of a listening environment, both in terms of its ability to support stereo imagery and its neutrality to these images. If a listening environment is unable to support the reproduction of strong spatial cues, then it is not very well adapted to audio reproduction. The kind of modifications that need to be made to listening environments are very selective; for example, since we know that the presence of early reflections in the listening environment can selectively distort or destroy certain spatial sound paths, it is not the entire range of reflected sound that needs control, but just the significant early reflections.

Our experience with this sound room has reinforced our awareness of the extent to which the unwanted effects of the listening environment are ignored in most sound reproduction. Clearly, the effects of early reflected sound can overwhelm the imagery produced by even the very best audio equipment. Virtually the entire audio community has come to accept the fact that sound imagery varies tremendously from room to room without questioning whether it really has to be tolerated. Many times the effects of early reflected sound are attributed to the audio equipment itself, especially the speakers, even though careful testing would reveal the true source of the problems. We feel that the listening environment should be considered as an extension of the speakers. Since the knowledge of how to address the inconsistency of audio reproduction exits, we believe the audio community should begin to insist on standards for listening environments. Beside the fact that sound imagery will be similar if not identical from room to room, it will be greatly improved in all rooms. Our simulation of directional cues based on idealized spectral transforms is improved in a controlled listening environment, and the advantage of having full three-dimensional sound imagery for stereo reproduction is quite clear.

C. The Number of Speakers

In order to provide listeners with aural experience of a complete three-dimensional space, appropriate auditory cues must be presented to both the left and right ears. This means that listening to spatially positioned images requires a minimum of two sound-producing transducers capable of creating different transformations of the original sound event at each ear. In the case that these transducers are headphones, the signal reaching each ear is completely predictable and the listener's experience is exactly that determined by the recorded material. In the case that these transducers are speakers, the signal reaching each ear is influenced by the listening room environment and there is a conflict between the spatial cues of the recorded material and of the listening environment. A number of researchers have been successful to varying extents applying knowledge of spectral cues to sound reproduction with speakers (Bloom, 1977; Mori, Fujiki, Takahashi, and Maruyama, 1979; Sakamoto, et al., 1981a, 1981b). Given the success of spectral cues even in somewhat hostile environments, how many speakers should be involved in the sound reproduction?

Multiple speaker placements have been suggested by many people. For example, Chowning, et al. (1974) proposed that the listener should be placed at the center of an imaginary tetrahedron with a speaker placed at each vertex. Their interest in this sort of speaker arrangement resulted from their recognition that the simulation of natural-sounding reverberation must include indirect sound arriving from elevated angles (their discussion emphasized the importance of the ceiling reflections in a cathedral). However evenly such a speaker arrangement seems to cover the possible angles from which sound events might emanate, it is not based on the concept of recreating the entire binaural sound field in which localization judgements are normally made. Panning amplitude between a speaker at ear level and an elevated speaker will not result in the same illusion of smooth spatial motion that occurs between speakers located on the left and right of the observer. More speakers do provide more potential sound source directions from within the listening environment, but will not provide the correct spatio-temporal distribution of reflected sound for an arbitrary illusory environment. Then too, if sounds are to be localized at positions in space where there are no speakers, then all the unwanted reflections in the playback environment must be greatly attenuated because they provide potential cues for localizing sound at speaker locations. If the real reflections are not overwhelmed by simulated reflections, then not only will sound sources be localized at the speaker's actual distance from the listener, but also the illusory angle of incidence may be rejected in favor of the speaker's angle of incidence.

Taken together, these considerations point to the conclusion that reproduction employing many speaker locations creates more problems for sound imagery and localization than it might seem to solve. Since sound delivery free from significant early reflections is desired, the smaller the number of speakers used, the simpler the task of optimizing the playback environment. And since illusions of all spatial locations can be supported by two transducers, stereo reproduction in a controlled acoustic environment is probably quite adequate for most purposes. It should be noted, however, that previous attempts at providing systematic transformations for placing sounds in three-dimensional space led Mori, et al. (1979) to conclude that a quadraphonic reproduction is required for creating the best distinctiveness between illusions of frontward and rearward locations. In typical "live" listening environments, the production of illusory distinctions between front and back locations is much more difficult than the production of distinctions between above and below locations. In fact, it is our experience that above/below distinctions can be supported under conditions of stereo reproduction in fairly "hostile" environments (those in which strong early reflections are uncontrolled). On the other hand, forcing sound delivered from in front of the listener to be heard as coming from directly behind the listener is practically impossible in such reverberant environments. This assertion leads to the conclusion that in uncontrolled reverberant environments, front/back distinctions require speakers in front and in back of the listener.

IV. SIMULATING REFLECTED SOUND

Our original localization experiments were based almost exclusively on the use of idealized pinna transfer functions and involved only one aspect of total spatial hearing: directionality. But, as has been discussed, superimposing pinna cues on a sound source does not always provide accurate directional sound images. Given the influence of reflected sound on perceived direction, one must construct some strategy for overcoming or controlling the reflected sound present in the listening environment.

Our approach to this problem is to simulate reflected sound along with pinna cues in such a way as to overwhelm the characteristics of the actual listening environment. For example, a sound source to be localized in front of the listener should have simulated early reflections arriving from many directions including those from behind the listener. The juxtaposition of front-localized direct sound and rear-localized copies of that direct sound arriving shortly after greatly increase the listener's ability to make front/back discriminations even in highly reverberant environments. Furthermore, even though most directional cues do not require support from simulated reflected sound, the apparent distance of sound sources at their particular angles of incidence does depend strongly upon the spatio-temporal distribution of the simulated reverberation (the distribution of reflections in space and time). Speakers need not be placed at multiple distances from the listener, since the relevant spatial and temporal patterns of indirect sound can be recreated exactly by two speakers. Reflected sound is simulated, then, for at least two reasons - to support directional cues and to place sounds at a given apparent distance.

A. Previous Approaches to Reverberation

The computer simulation of reflected sound has been a topic of interest in computer music and in audio research ever since the publication of Manfred Schroeder's (1962; 1970) pioneering articles. The technique of recirculating delays described by Schroeder in "Natural Sounding Artificial Reverberation" has served as the basis for reverberation simulation in computer music ever since its publication in 1962. More recent refinements in the art of simulating reflected sound have mainly modified Schroeder's scheme in order to achieve some higher degree of naturalness and realism. Chowning's (1970) scheme for simulating the movement of sound sources in a reverberant space manipulated the balance between direct and reflected sound in order to provide the listener with realistic cues to the perceived distance of a sound source. Moorer's (1979) redesign of the signal processing elements used in the Schroeder reverberator simulates the way in which natural reverberation becomes increasingly low-pass with time as high frequencies in the sound are absorbed.

These and other contemporary approaches growing out of the Schroeder reverberator share a number of basic assumptions about the particular quality of the reverberation they intend to produce and about the way in which the sound will be delivered to the listener. First, they all attempt to replicate the kind of global reverberation that is typical of large reverberant rooms such as concert halls. Even though such reverberation is appropriate for many kinds of music, it is not appropriate for the full range of possible applications of spatializing sound with a computer. Second, these schemes attempt to capture the general characteristics of reverberation in large rooms without attempting to replicate any of the exact characteristics that distinguish one room from another. Neither do they attempt to make provisions for changes in the location of the listener or the size and shape of the reverberant environment. In effect, since these methods do not actually model a room, controlling the characteristics of the reverberation is largely a matter of guess work. Third, all these methods are intended for use in conventional stereo (or quadraphonic) reproduction and make no attempt to localize or spatially separate the reverberant sound. The one exception to this last point is the reverberator designed by Stautner and Puckette (1982). Their reverberator attempts to capture the distribution of reflected sound in real rooms by providing each output channel with reverberation that is statistically similar to that coming from part of a reverberant room.

B. Spatial Reverberation

Our first attempt to add reverberation to sounds processed with directional hearing cues created a very unnatural situation. The spatialized sound sources moved out away from the speakers in the intended path, but the reverberation remained at the speakers. Not only was the reverberation spatially static, but its subjective character also did not change as the sound source moved. We reasoned that a more natural reverberant field could be created by adding simulated early reflections to the direct signal and its reverberation. After specifying the dimensions of a model room, the sound source path and listener position, we calculated the time-delay and direction of first-order reflections reaching the listener (those having been reflected from a single wall). We simulated not just changes in the time-delay of each "first-order" reflection, but also changes in each reflection's direction as the sound source moves about the room. Listening tests demonstrated that the spatial quality of the direct sound had improved, but that the reverberation stayed at the speaker positions. Subsequent experiments with adding second- and third- order reflections showed improvement but did not overcome the problem of creating a truly natural-sounding acoustic field.

Our particular experience with reflected sound and pinna cues has led us to the initial design of a reverberator with different assumptions and goals than those discussed earlier. We have concluded that in order to simulate the spatial cues of real environments, one must capture the total spatio-temporal pattern of reflected sound. For this reason, we have sought a reverberator design that models an actual room and which accurately replicates the spatial and temporal distribution of reflected sound. The design must differentiate between large and small rooms and allow us to place the reverberated sound source anywhere in three-dimensional space, not just at the speaker positions. A basic signal processing network for spatial reverberation requires two subsystems (Figure 5). One is a reverberation subsystem that takes an input signal and produces multiple outputs, each of which is a unique "reverberation stream". The temporal pattern of each stream must match the reflections from a small spatial region of the model room. The other is a "directionalizer" subsystem that superimposes directional hearing cues such as pinna filtering. Each reverberation stream is individually directionalized to position the simulated reflections in the correct region of the model room. The sum of all reverberation streams taken together captures the entire spatio-temporal pattern of reflected sound in the model room (even though simulated reflections are spatially bunched together).

We believe that this spatio-temporal pattern creates the context in which judgements of direction and distance are made. A mental model of the acoustic environment develops rapidly from this context during normal experience in a novel environment. If one is denied normal exposure to an environment, such mental models are not formed and localization ability is significantly degraded (Musicant and Butler, 1980). We also believe that the spatio-temporal pattern formed by the reflected sound can serve to clarify the position of the sound source in the environment, especially once the sound source begins to move. The reflected sound is particularly important to directional judgements when other directional cues are weak, such as the cues for distinguishing between front and back positions. Thus, our primary goal is that the spatial reverberator produce the kind of spatially distributed reverberation that will help listeners localize sounds and impart a very strong impression of the reverberant environment.

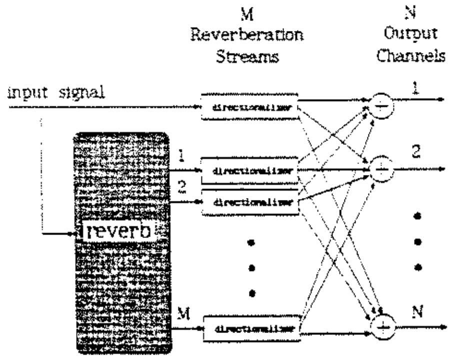

Figure 5 A reverberation network producing directionalized reverberation streams.

C. Implementing the Image Model

The concept of spatial reverberation is sufficiently broad to encompass a number of different implementations. The one we shall discuss here is based on the most commonly discussed model of reflected sound in real environments called the "image model" (Allen and Berkeley, 1979). Under this model, reflected sound is viewed as emanating from "virtual rooms" outside of a physical room. Each virtual room is a normal-image or mirror-image replication of the physical room that contains a copy of the sound source called a "virtual source." The total pattern of reflected sound in the physical room can be viewed as the composite sound reaching the listener from all virtual sources (illustrated in Figure 6). Drawing lines between these virtual sources and the position of a listener in the physical room will predict the points along the wall from which the reflected sound emanates. Even though true wall reflections involve spatial diffusion, this model provides a good approximation of the direction and timing of the most important reflections.

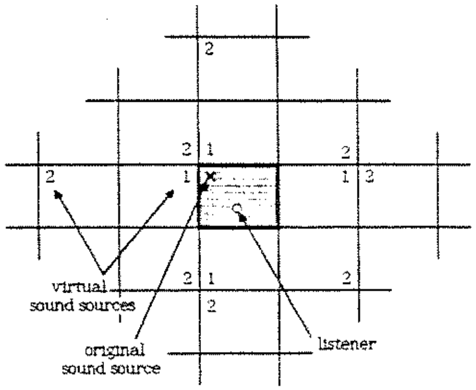

Figure 6 Top view of virtual image rooms in the horizontal plane

Figure 7 illustrates a listener's perspective on the spatial distribution of the virtual sources for first-, second- and third-order reflections as predicted by the image model. The center of this plot is the point directly in front of the listener at ear level. The position of each circle represents the incidence angle of a discrete early reflection. The horizontal dimension is azimuth angle, ranging from -180 to 180 degrees (circles located at both the extreme left and the right of the plot indicate reflections arriving from behind the listener). The vertical dimension is elevation, extending 90 degrees above and below ear level. In this example, a sound source (represented by the asterisk) is located at an azimuth angle of 90 degrees (to the right of the listener), an elevation angle of 30 degrees, and a distance of 1.5 meters. The modeled room has length, width, and height of 4.5, 5.5, and 6 meters, respectively. The listener is seated just behind the center of the room. The time delay on each reflection relative to the arrival of the original sound source is indicated by the size of the circle – circle size decreases exponentially with time delay.

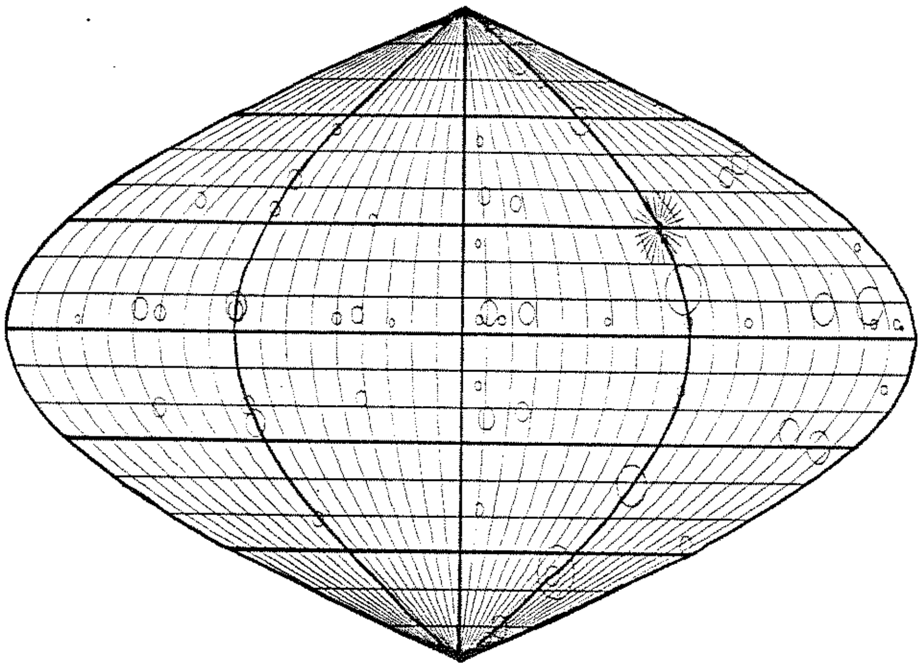

Figure 7 Spatio-temporal distribution of early reflections Horizontal axis is azimuth angle, vertical axis is elevation angle, circle size represents reflection latency (size diminishes exponentially with time).

We have developed a signal-processing network that closely matches the spatio-temporal distribution of reflected sound predicted by the image model. This signal processing network represented in Figure 8 is an elaboration of the general approach shown in Figure 5. There are still separate directionalizers for each reverberation stream, but the reverberation network takes on a form designed to match image model predictions. Values for time-delay, amplitude and direction are calculated from user specified information on the size of the simulated room, the location and orientation of the listener and the location of the sound source.

As in the earlier version, the input signal passes both into a directionalizer and into the reverberation network. The signal enters the network through a non-recirculating delay buffer with multiple taps. There are six first-order reflections predicted by the image model that emanate from virtual rooms behind the six walls of the simulated room. Six delay taps are used to replicate the time-delays of these first-order reflections while the gain of each reflection is produced by multiplication with the scaling coefficients, a[1-6].

The signal from each of these delay taps is passed to the inner reverberation network and is also separately directionalized to produce a discrete simulated reflection. The image model predicts a total of eighteen second-order reflections. Six of these second-order reflections originate in virtual rooms directly behind the first-order virtual rooms. The six first-order delays from the delay buffer are fed into the inner reverberation network where they are used to generate reverberation streams that start with these particular second-order reflections. The remaining twelve second-order delays originate in second-order virtual rooms that touch the junction of two walls in the model room. Twelve taps from the delay buffer are used to replicate the time-delays for these reflections. They are passed directly into the inner reverberation network where they are used to generate reverberation streams that begin with these second-order reflections. The spatial reverberator provides separate directionalization for all of the eighteen reverberation streams that begin with second-order reflections. The output of each directionalizer represented in Figure 8 has both left and right channels which must finally be mixed to provide two channels of stereo output.

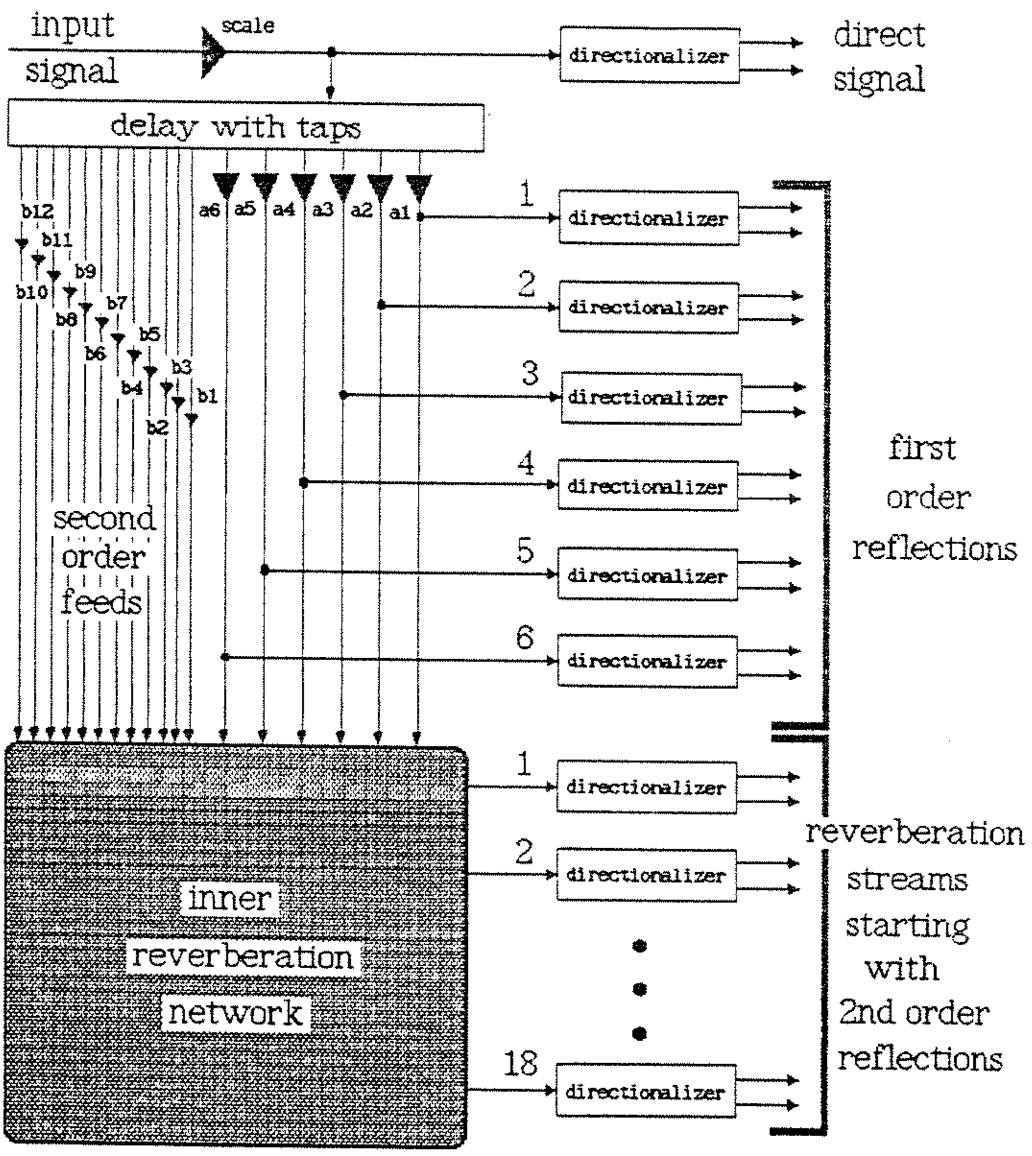

Figure 8 The image model implementation of the spatial reverberator

Describing the inner reverberation network is sufficiently complex as to be beyond the scope of this paper (but a subsequent paper by Kendall and Martens will focus on the signal processing attributes of this system). For the purposes of this discussion we need only mention that all reflections predicted by the image model are accounted for. Time delays are exact for all second-order and most third-order reflections. Higher-order reflections are statisically correct in time; the higher the order, the greater the possible error. All second-order reflections have exact directions and higher-order reflections are localized in the same direction as a nearby second-order reflection. Although, individual reflections may sometimes be displaced by 30-degrees, the overall spatio-temporal pattern of reflections in the model room is clearly captured. The reflection patterns of large rooms are quite distinct from those of small rooms, as are long rooms from short rooms.

The particular implementation we have discussed here has the virtue of matching a well-understood model of room reverberation. It is important that the implementation of a spatial reverberator at least be able to match the complexity of the image model, but it is not so important that it match the exact solution. The same signal-processing network that we have described here can be used to generate sound fields quite different from those predicted by the image model. Changes in the time delay and spatial position of reflected sound can radically alter the perceived sound space. In fact, one can ignore physical models of reverberation and select the spatio-temporal pattern of reflected sound from entirely different perspectives. For instance, one can choose a kind of spatial reverberation that ignores the physical simulation in favor of manipulating subjective characteristics.

V. USER INTERFACE

TO THE SPATIAL REVERBERATOR

A. The Issue of Control Parameters

The processing required by the spatial reverberator may seem somewhat complex. The question must be raised, "What does all this complexity buy, that a less complex approach cannot accomplish?" The answer is that it gives us the potential for unprecedented control over the subjective characteristics of the room reverberation. Though it is certainly the case that the realistic binaural sound field created by spatial reverberation allows us to create very vivid spatial perceptions of sound sources, there are many advantages to accurate simulation of reverberation other than its ability to support accurate localization. Spatial reverberation can potentially create spatial percepts of all sorts. Our spatial reverberator was initially designed to accept control parameters that specify the physical attributes of a room – such as the room's dimensions, the absorption coefficient of the walls, the physical positions of the listener and sound source, etc. But composers or others who wish to use the reverberator probably will want to specify control parameters that have more psychologically relevant meaning.

Even though reflected sound may be considered a two-dimensional physical phenomenon (the indirect sound field having spatial and temporal characteristics), there are a numerous psychological dimensions associated with it. Besides its role in perceived direction and distance, reflected sound plays a primary role in subjective qualities related to the acoustics of the room. In the actual spatial experience of real rooms, the perception of direction and distance are inseparably linked to the subjective room acoustics. All these factors interact in the formation of our subjective impression of the acoustic environment and the position and quality of a sound within that environment. It is easy to overlook the importance of subjective room acoustics because our everyday working notion of room quality is less precise than our notion of direction and distance. Then too, we rarely have experience with unnatural environments that would bring these qualities out of the perceptual background. Nonetheless, they play an important role in total spatial hearing and are especially important to us because of their aesthetic relationship to the performance of music.

We view a large part of the task of perfecting our spatial reverberator as learning how to create a user interface with intelligence about the relationship of these psychological dimensions to a room's physical attributes. This interface needs to take a list of psychological parameters which constitute a multidimensional psychological description and convert it into a set of physical simulation parameters – a relation for which there is not a one-to-one correspondence. Without an interface of this sort, composers would be forced to guess at the physical attributes of rooms which they imagine primarily in subjective terms.

B. The Subjective Dimensions of Sound Space

Although it is clear that reflected sound has a profound influence on the formation of sound images under normal listening conditions, it is not well understood what aspects of reflected sound influence what subjective qualities. Nonetheless, we can analyze our experience of sounds in space in terms of distinct perceptual dimensions. In the literature of subjective room acoustics, there is no common terminology for describing the subjective dimensions of sound space (for review, see Rasch and Plomp, 1982). Besides the directional aspects of sound in space, there are four subjective dimensions which we believe are important to the manipulation of spatial imagery in music production and reproduction:

a) **Distance quality** – the immediate percept of egocentric distance, rather than that which may be inferred from the ratio of indirect to direct sound

b) **Definition** – perception of sound-source characteristics; eg., spatial extent or focus

c) **Spatiousness** – perception of environmental characteristics; eg., liveness, size, and shape

d) **Spatial texture** – the perception of the interaction of the sound with its environment

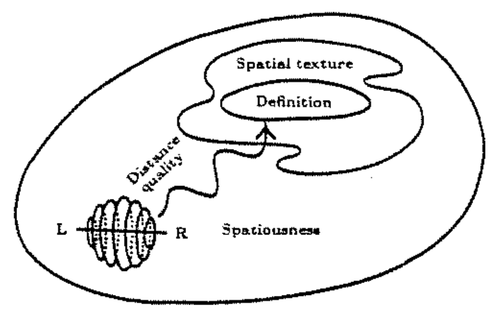

Figure 9 Subjective dimensions of sound space: In addition to the directionality of a sound in space, there are the characteristics of distance, definition, spatiousness, and spatial texture

Distance quality. It is generally accepted that distance is judged largely on the basis of the loudness of the direct sound and its relation to the loudness of the reverberation. But under natural listening conditions, it is the exact spatio-temporal pattern of reflected sound that plays a dominant role in the perception of distance. For instance, Mershon and King (1975) argue that natural reverberation provides an absolute cue to distance while changes in loudness provide only relative cues to changes in perceived distance (cf. Sheeline, 1982). Current methods for creating the illusion of distance rely only upon the relative proportion of direct energy to reflected energy. Chowning's (1971) four-channel reverberation scheme illustrates the culmination of this approach which, though quite effective in simulating distance cues for large environments, cannot provide highly salient cues to distance for smaller environments.

Our own experience with the simulation of cues to distance using the spatial reverberator has revealed some unexpected facts about the minimal conditions for creating illusions of sound source distance. In the early stages of developing the concept of spatial reverberation, we used convolution techniques to generate test stimuli which contained directionalized direct signal and directionalized reflections within a short time window. These test stimuli were generated for numerous spatial paths that varied in azimuth, elevation and distance. The illusions of changing distance of the sound source were quite powerful even when the only reflections simulated were those which followed the direct signal by 33 msec or less. The distance of the sound source was perceived immediately, though the spatial image was quite "dry" and gave little impression of the model room upon which the simulation was based. We are forced to conclude that the early reflections experienced in typical (even very small) listening environments provide completely sufficient information for localizing sounds at a particular distance from the listener.

Definition and spatiousness. These two subjective properties of room acoustics are often described as complementary to one another; that is, a high degree of definition precludes a high degree of spatiousness (Rasch & Plomp, 1982). Definition is generally defined as the property of a sound object having to do with the width or focus of its image (also referred to as "spatial extent" by Blauert, 1982). Sound images which are highly defined in their spatial locations are usually highly intelligible, but it should be stressed that intelligibility always is dependent upon temporal characteristics of the sound objects and reverberation as well as the spatial separation of the sound objects (Haas, 1972; Santon, 1976). Spatiousness, on the other hand, is defined as a property of the sound environment itself. There are many terms which seem to be used somewhat interchangeably with this term (eg., "ambience," "presence," "reverberance," etc . . .). but all refer to the perceived size of the space and how live or reverberant that space seems. Concert halls are generally designed to maximize spatiousness at the expense of good definition. In the most highly rated concert halls, sound seems to reach the listener from all directions regardless of the actual source location (Ando, 1983) and thus provides many listeners with similar sound images regardless of where they sit. Borish (1984) reported that the concert hall geometry that provides such a reverberation pattern is the "shoe box" shape rather than the "fan" shape. The loss of spatial definition that occurs when reverberation is independent of source location may seem unavoidable when highly spatious imagery is desired, but definition and spatiousness are influenced by other factors as well.

Kurozumi and Ohgushi (1983) have shown that perceived spatiousness is predicted by the interaural cross-correlation (IACC) of the binaural sound. Most reverberant environments produce acoustic signals at the two ears which have low IACC and are perceived as highly spatious. Environments with little reverberation can produce acoustic signals at the two ears that are more highly correlated and potentially perceived as high in definition. But it appears that not all reflected sound contributes to spatiousness in the same way: reflected sound has differing impact on spatiousness and definition depending on its spatial distribution. Gottlob (1975) found that while lateral reflections are necessary for producing highly spatious imagery, definition was degraded by lateral reflections. When only non-lateral reflections were present, perceived definition was improved. A more objective confirmation of this phenomenon is Hartmann's (1983) finding that localization accuracy (a measure of positional definition in a spatial image) for impulsive sounds was affected by the direction from which early reflections appeared. When early reflections came from the ceiling and floor rather than the lateral walls, subjects showed fewer errors in localization.

Given this knowledge it is clear that spatiousness and definition can be controlled in part with a spatial reverberator. If one chooses to simulate a very reverberant room with dominant early reflections from the lateral walls, then highly spatious images will be created at the expense of good definition. If the emphasis is given to the ceiling reflections, then high definition can be reinforced. We are currently experimenting with strategies for creating high definition and good spatiousness at the same time.

There are other ways in which controlling the entire spatio-temporal pattern of reflected sound provides superior illusions of the sound environment. We are experimenting with the ability of listeners to hear relatively subtle changes in the rooms simulated by the spatial reverberator. For example, when the ratio of direct sound to indirect sound is varied while holding global reverberation time fixed, the number of distinguishable spatiousness percepts that can be obtained is quite limited (Reichardt and Schmidt, 1966). Informally repeating this experiment with spatial reverberation rather than global reverberation shows that for a fixed reverberation time, the spatiousness may be experienced as changing in room liveness or changing in room size. Furthermore, spatial reverberation can create distinctions between long hallways and deep wells of the same size and shape (these having identical temporal distributions of reflections).

Spatial Texture. Our experience with the variety of spatial images that spatial reverberation can create has led us to expand upon the concepts found in the subjective room acoustics literature. In addition to the perceptual dimensions of definition and spatiousness described by Rasch and Plomp (1982), we have identified a third which we term "spatial texture." We define spatial texture to be the quality of the sound object in its environment. Although this property is a function of environmental reflected sound, it is perceived as a quality of the object. It is the particular quality imbued to a sound object due to its position in the environment, a quality that changes with different locations and different kinds of rooms. Though it is difficult to explain the concept of spatial texture when few people have consciously recognised it, we can give it the kind of negative definition often given to timbre: Whereas timbre has been described as the quality which differs between two tones of the same pitch and loudness, spatial texture may be described as the quality which differs between two spatial images having the same distance, definition, and spaciousness. As is the case with most other subjective qualities of the spatial image, spatial texture is the product of an interaction of all spatial and temporal information associated with a sound event. Nonetheless, we can present an example of what features of reflected sound seem to produce differences in spatial texture.

Figure 10 illustrates a spatial path for a sound source which maintains a constant distance from the listener. Let's assume that we can adjust indirect sound characteristics as the sound source moves along this path so as to maintain constant apparent distance and definition of the sound source, and constant perceived spatiousness of the space through which the sound is passing. As the sound source moves from in front (Figure 10a), to 22 degrees to the right (Figure 10b), to 45 degrees (Figure 10c) what changes is the spatio-temporal distribution of the early reflections. This creates a distinct spatial impression for each of these three sources.

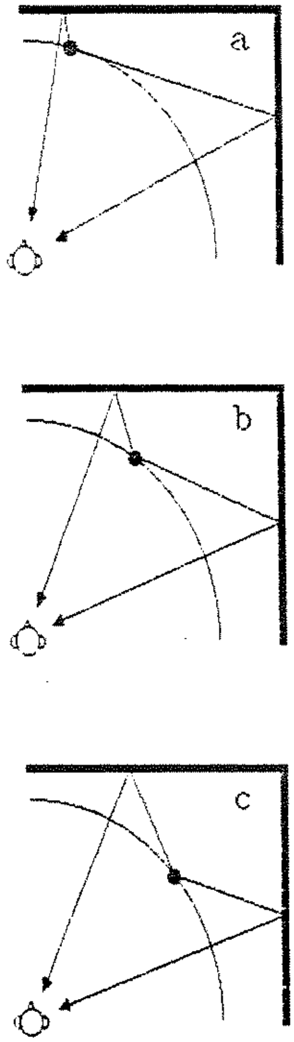

Figure 10 Changes in the pattern of early reflections associated with changes in spatial texture.

C. Application to Sound Design

The psychological dimensions of sound space are manipulated in various ways in today's recording studio. The microphone placement, the reverberant sound in the recording environment, the use of artifical reverberation, and the particular mix of these for different tracks all affect psychological dimensions associated with sound space. The technique of spatial reverberation permits direct manipulation of these psychological qualities as well as of perceived location. When separate tracks of processed music are combined, the spatial reverberator has the power to control the degree of blend and/or separation of the individual parts. It also has the power to affect immediacy and/or warmth of the sound objects. In fact, the psychological dimensions of sound space may provide a powerful metaphor to the kinds of qualitative manipulations performed by audio professionals.

It was initially very important to us that spatial reverberation be able to model a physical room and create a realistic sound field. For the purpose of musical applications in sound design the link to physical modeling is clearly not important. We have already stated that creative professionals using spatial reverberation will most likely communicate their intentions in psychological terms rather than in physical models. It is also the case that providing independent controls of the psychological dimensions of sound space will subvert the physical model on which spatial reverberation was originally based. For example, "definition" and "spatiousness" may not be independent of each other in actual physical rooms, but under computer control the independence of these qualities may be achieved with spatio-temporal patterns of reflected sound that do not fit a room that can be physically realized. We have designed a plan for the systematic evaluation of the spatial imagery that our reverberator can produce. The first stages of this research involve collecting similarity judgments for musical material processed by the spatial reverberator and submitting these data to multi-dimensional scaling analyses as in Sheeline's (1983) work. Although a totally comprehensive evaluation of perceptual attributes of the spatial reverberator may require years of effort, we expect that the foundational work will be completed in a relatively short time.

How spatial reverberation is used to process music may evolve not as a technical issue, but as a stylistic and aesthetic one. It remains to be seen how spatial reverberation is best applied to the audio processing of different kinds of music. It certainly is not necessary that an orchestral recording, for example, recreate the experience of a live performance. Spatial reverberation provides an opportunity to reinterpret the spatial structure of preexistent music. More importantly, it provides composers with the means to create an original spatial structure for music conceived with spatial reverberation in mind.

VI. OBSERVATIONS AND PRESCRIPTIONS

The simulation of sound space has up to this time been approached in a piece-meal fashion with one or another relatively crude approximations to real spatial cues. We have all become terribly acclimated to the results, but only a little reflection is needed to realize that the limited success of these simulations is built on extremely deformed and impoverished spatial cues. Any serious attempt to simulate sound space must be built on the concept of total spatial hearing cues and a proper concept of the nature of sound space. This means synthesizing all directional cues and simulating reflected sound. This also means grasping the important perceptual dimensions of sound space and taking the qualities of definition, spatiousness and spatial texture into consideration along with direction and distance. We hope that an appreciation for the unique nature of sound space and its perceptual complexity will inspire composers to put space to new uses in their compositions and to integrate it into the structure of their works in new ways.

ACKNOWLEDGEMENTS

This work has been supported by a grant from the System Development Foundation. Special thanks must be given to Charles Smith whose commitment to new ideas and whose confidence in our abilities were crucial to the project. We would also like to thank Pablo Picasso for his line drawing of Igor Stravinsky which appears in Figure 1.

REFERENCES

Allen, J. B. & Berkeley, D. A. (1979) Image method for efficiently simulating small room acoustics. Journal of the Acoustical Society of America, 65, 943-950.

Ando, Y. (1983) Calculation of subjective preference at each seat in a concert hall. Journal of the Acoustical Society of America, 74, 873-87.

Barron, M. F. E. (1971) The subjective effects of first reflections in concert halls – the need for lateral reflections. Journal of Sound Vibration, 15, 475-94.

Batteau, D. W. (1967) The role of the pinna in human localization. Proceedings of the Royal Society of London, 168 (series B), 158-180.

Blauert, J. (1969) Sound localization in the median plane. Acustica, 22, 957-962.

Blauert, J. (1981) Lateralization of jittered tones. Journal of the Acoustical Society of America, 70, 694-698.

Blauert, Jens (1982) Spatial Hearing, trans. John S. Allen. MIT Press (Cambridge, Mass.) 1983. Originally entitled Raumliches Horen, S. Hirzel Verlag, Stuttgart, 1974.

Bloom, P. J. (1977) Creating source elevation illusions by spectral manipulation. Journal of the Audio Engineering Society, 25, 560-565.

Borish, J. (1984) Extension of the image model to arbitrary polyhedra. Journal of the Acoustical Society of America, 75, 1827-1839.

Butler, R. A., & Belendiuk, K. (1977) Spectral cues utilized in the localization of sound in the median sagittal plane. Journal of the Acoustical Society of America, 61, 1264-1269.

Chowning, J. M. (1971) The Simulation of Moving Sound Sources. Journal of the Audio Engineering Society, 19, 2-6.

Chowning, John. M., Grey, John M., Moorer, James A. & Rush, Loren (1974) Computer Simulation of Music Instrument Tones in Reverberant Environments. Department of Music, Stanford University, Technical Report STAN-M-1.

Davis, D. & Davis C. (1980) The LEDE Concept for the Control of Acoustic and Psychoacoustic Parameters in Recording Control Rooms. Journal of the Audio Engineering Society, 28, 585-95.

Feddersen, W. E., Sandel, T. T., Teas, D. C., & Jeffress, L. A. (1957) Localization of high-frequency tones. Journal of the Acoustical Society of America, 29, 988-991.

Gottlob, D. (1975) Vergleich objektiver akustischer Parameter mit Ergebnissen subjektiver Untersuchungen an Konzertsalen. PhD Dissertation, Georg-August-Universitat, Gottingen.

Hartmann, W. M. (1983) Localization of sound in rooms. Journal of the Acoustical Society of America, 74, 1380-1391.

Haas, H. (1972) The influence of a single echo on the audibility of speech. Journal of the Audio Engineering Society, 20, 145-159.

Kendall, G. S., & Rodgers, C. A. P. (1982) The simulation of three- dimensional localization cues for headphone listening. Procedings of the 1982 International Computer Music Conference.

Kurozumi, K. & Ohgushi, K. (1983) The relationship between the cross-correlation coefficient of two-channel acoustic signals and sound image quality. Journal of the Acoustical Society of America, 74, 1726-33.

Mershon, D. H. & King, L. E. (1975) Intensity and reverberation as factors in auditory perception of egocentric distance. Perception and Psychophysics, 18, 409-415.

Moorer, J. A. (1979) About This Reverberation Business. Computer Music Journal, 3, 13-28.

Mori, T. Fujiki, G., Takahashi, N., Maruyama, F. (1979) Precision sound-image-localization technique utilizing multitrack tape masters. Journal of the Audio Engineering Society, 27, 32-38.

Morimoto, & Ando, Y. (1979) Simulation of Sound Localiztion. Presented to the Symposium of Sound Localization, Guelph, Canada.

Musicant, A. D. & Butler, R. A. (1980) Monaural Localization: An Analysis of practice effects. Perception and Psychophysics, 28, 236-240.

Nuetzel, J. M., & Hafter, E. R. (1976) Lateralization of complex waveforms: Effects of fine structure, amplitude, and duration. Journal of the Acoustical Society of America, 60, 1339-1346.

Rasch, R. A. & Plomp, R. (1982) The Listener and the Acoustic Environment. In: D. Deutsch (Ed.) The Psychology of Music. New York: Academic Press.135-147.

Reichardt, and Schmidt (1966) Die horbaren Stufen des Raumeindrucks bei Musik. Acustica, 17, 175-9.

Rodgers, C. A. P. (1981) Pinna transformations and sound reproduction. Journal of the Audio Engineering Society, 29, 226-234.

Santon, F. (1976) Numerical prediction of echograms and of the intelligibilty of speech in rooms. Journal of the Acoustical Society of America, 59, 1399-1405.

Sakamoto, N., Gotoh, T., Kimura, Y. (1976) On "out-of-head localization" in headphone listening. Journal of the Audio Engineering Society, 24, 710-715.

Sakamoto, N., Gotoh, T., Kogure, T., Shimbo, M. (1981a) Controlling sound-image localization in stereophonic reproduction: Part I. Journal of the Audio Engineering Society, 29, 794-799.

Sakamoto, N., Gotoh, T., Kogure, T., Shimbo, M. (1981b) Controlling sound-image localization in stereophonic reproduction: Part II. Journal of the Audio Engineering Society, 30, 719-721.

Schroeder, M. R. (1962) Natural-Sounding Artifical Reverberation. Journal of the Audio Engineering Society, 10, 219-23.

Schroeder, M. R. (1970) Digital Simulation of Sound Transmission in Reverberant Spaces. Journal of the Acoustical Society of America, 47, 424-431.

Shaw, E. A. G. & Teranishi, R. (1968) Sound pressure generated in an external-ear replica and real human ears by a nearby point-source. Journal of the Acoustical Society of America, 44, 240-9.

Sheeline, C. W. (1982) An Investigation of the Effects of Direct and Reverebrant Signal Interaction on Auditory Distance Perception. PhD dissertation. Department of Hearing and Speech Sciences, Stanford University.

Stautner, J. & Puckette, M. (1982) Designing multichannel reverberators. Computer Music Journal, 6, 52-65.

Wright, D., Hebrank, J. H., & Wilson, B. (1974) Pinna reflections as cues for localization. Journal of the Acoustical Society of Amerfca, 56, 957-62.