7 Spatial Reverberation: Discussion and Demonstration

Gary S. Kendall, William L. Martens, and Shawn L. Decker

7.1 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. Multichannel 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 that is spatially impoverished. This is despite the fact that many 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 that have no spatial attributes save those provided by signal processing.

The goal of the work described here is to provide composers with a comprehensive control of auditory space percepts in music. 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.

7.2 Directional Hearing Cues

7.2.1 Interaural Differences

The two cues that 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 microsec (Feddersen et al. [8]). 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 a salient cue only 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 ITDs is abolished if the high-frequency stimuli have time-varying amplitude (e.g., Nuetzel and Hafter [12]) or time-varying frequency (e.g., Blauert [4]).

In actuality, IID and ITD provide the auditory system only 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 et al. [14]). 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 1960s. 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 perpendicularly 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 7.1. This circle is a representation of the possible locations in three-dimensional space at which the lateralized sound might have originated.

7.2.2 Spectral Cues

Within the last 20 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), the 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 an asymmetric

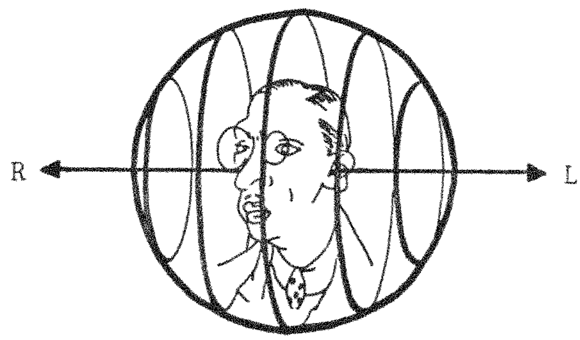


Figure 7.1
The left/right lateralization dimension and the up/down/front/back circles upon which sounds can be located at a given distance from the listener.

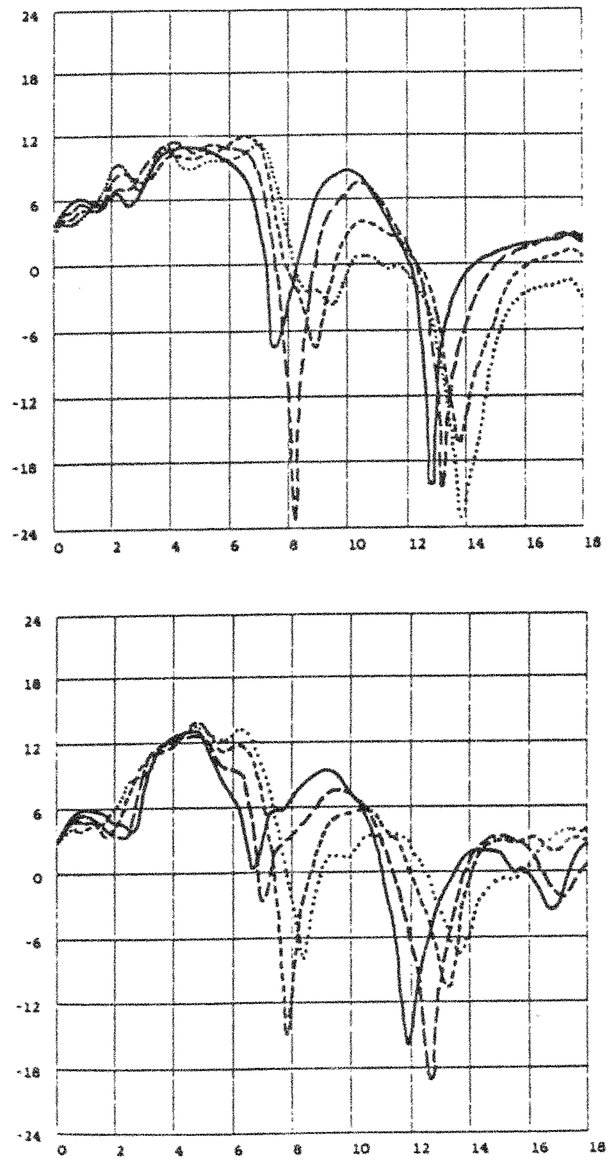
arrangement of ridges, the composite sound reflections create a unique spectral profile for every sound direction. The auditory system uses these spectral profiles to remove the spatial ambiguity that results from IID and ITD cues alone. It is the spectral cues produced by the pinna, shoulders, and upper torso that enable the auditory system to determine the position of the sound source on the circle represented in figure 7.1: 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) [2] 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 judgments of azimuth and for elevation. Shaw and Teranishi (1968) [17] 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) [3] 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 bridged the time-domain/frequency-domain distinction and demonstrated the essential similarity of both views.

Despite the fact that pinna transfer functions are highly complex and difficult to adapt 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 in spite of 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 7.2A) and subject GSK (figure 7.2B), 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 characterized largely 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 [6] 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.

7.2.3 Simulating Cues for Directional Hearing

For the purposes of simulating cues for directional hearing in computer music or any kind of audio reproduction, one must 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 estab-



(A)

(B)

Figure 7.2
Head-related transfer functions for two subjects (MDL in 7.2A and GSK in 7.2B) 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°) to an elevation 30° above ear level (solid line, 0°; long dashes, 10°, short dashes, 20°; dotted line 30°). Abscissa is frequency in kHz, ordinate is gain in dB.

lished 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 cannot be determined by averaging techniques because directional judgments 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 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 [6] 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 that 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° apart in azimuth and 20° apart in elevation. We have analyzed our empirical data in order to identify those spectral features that seem to be 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 pinnae, a set of spectral manipulations can be devised that not only match the saliency of natural cues but actually support superior directional perception for many people.

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. Figure 7.3A shows transfer functions measured inside the ear canal for two subjects. Figure 7.3B 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 notice-

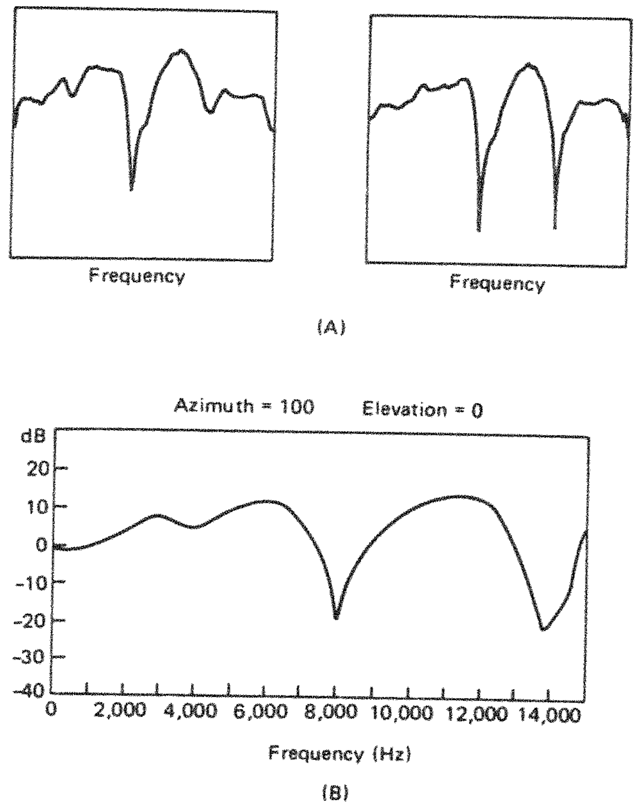


Figure 7.3
Comparison of measured head-related transfer functions for two subjects (upper panels: maximum is 18 kHz) and our simulation (lower panel: maximum is 15 kHz).

able noise. In effect, the filters must be designed for the entire ensemble of directions taken together.

7.3 Simulating Reflected Sound

7.3.1 The Spatial Reverberation Concept

Our particular experience with pinna cues led us to the conclusion that in order to simulate the spatial cues of real environments, one must capture the total spatiotemporal pattern of reflected sound. For this reason, we have sought a reverberator design that models an actual room and that

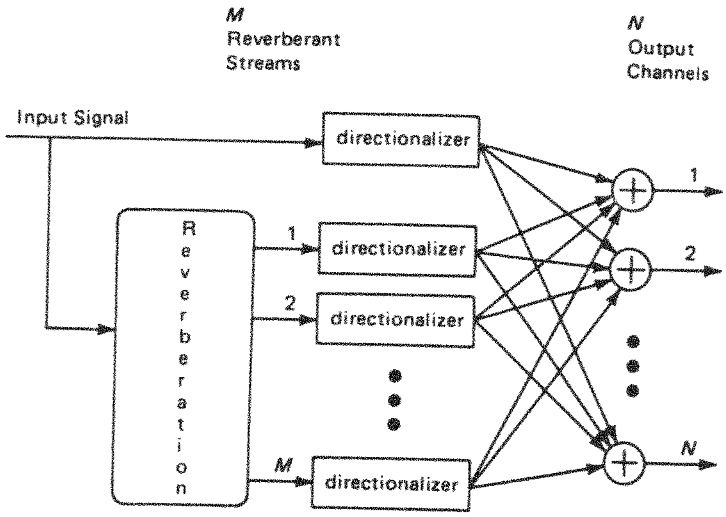


Figure 7.4
Basic signal-processing network for spatial reverberation.

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 7.4). 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 in 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 spatiotemporal pattern of reflected sound in the model room.

We believe that this spatiotemporal pattern creates the context in which judgments 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 accuracy is significantly degraded (Musicant and Butler [11]). We also believe that the spatiotemporal 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 judgments when other directional cues, such as the cues for distinguishing between front and back positions, are weak. 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.

7.3.2 Implementing the Image Model with Recirculating Delays

The technique of recirculating delays has served as the basis for reverberation simulation in computer music and digital audio production ever since the publication of Manfred Schroeder's pioneering work in 1962 [15]. The Schroeder reverberator prescribes small units of recirculating delays combined in a predominantly serial network. The complete reverberation network suggested by Schroeder consists of four comb filters in parallel followed by two all-pass filters in series. Although the comb filters produce a series of spectral notches in the reverberated signal, the four filters in parallel produce a density of notches akin to that measured in real rooms.

In 1970, Schroeder expanded his reverberation network to include an initial delay buffer that replicates the kind of the early reflection pattern typical of concert halls [16]. This model, including both the frequency-dependent recirculating delays and the simulation of early reflections, was refined by Moorer [10], who produced the best-sounding reverberation network of this type. A significant advance in the goals of such networks was achieved by Chowning [7], who combined the basic serial network of recirculating delays with time-variant controls to simulate moving sound sources.

These and other contemporary approaches to reverberation attempt to replicate the global reverberation typical of large reverberant rooms like concert halls without attempting to capture any of the exact characteristics that distinguish one room from another. Since these methods do not actually model a room, matching the reverberation to the characteristics of a particular room is largely a matter of guesswork. Even Chowning's system for simulating moving sound sources does not change the pattern of reflected sound in a way that captures the changes typical of real rooms.

For simple rooms the "image model" of reverberation (Allen and Berkeley [1]) provides a method for predicting both the spatial and the temporal pattern of reflected sound from the room dimensions and the positions of the sound source and the listener within that room. For this model, each ray of reflected sound is viewed as originating from a "virtual sound source" outside of the actual physical room. Each virtual sound source is contained within a "virtual room" that replicates the physical room or is a mirror image of it. We use three-dimensional coordinates (x, y, z) to specify individual virtual sound sources or virtual rooms (figure 7.5). The pattern of reflected sound in the physical room can be viewed as the composite sound reaching the listener from all virtual sources. A vector connecting the position of the listener in the physical room and these virtual sources will predict the direction from which the reflected sound emanates and the distance that the sound must travel before reaching the listener. Even though real walls usually have irregularities that cause reflections to be spatially diffused, the image model provides a good approximation of the direction and timing of the most important reflections.

Until now, the image model was computationally too expensive for use in audio and music production since it could only be implemented by convolution with an entire simulated room impulse response. It has also been generally believed that the image model could not be captured with recirculating delays. The reverberation network described in the following section is able to capture image model reverberation with recirculating delays.

Delays for the Source and First-Order and Second-Order Reflections. This basic signal-processing network can be realized in a number of ways. Figure 7.6 illustrates the version of the system implemented by the *space18* program currently in use at Northwestern Computer Music. The input signal to *space18* is passed into three different nonrecirculating delay buffers. Digital signal interpolation must be performed on all buffers when the delay times change (Smith [19]). The first buffer captures delays for the source itself. Its input is scaled and filtered to capture intensity and spectral changes due to distance; its output is passed directly to a directionalizer. The source signal enters the inner reverberation network through two delay buffers with multiple taps. Each buffer is preceded by a filter that captures the spectral changes due to air and wall absorption. The first buffer produces the delays for the six first-order reflections predicted by the image

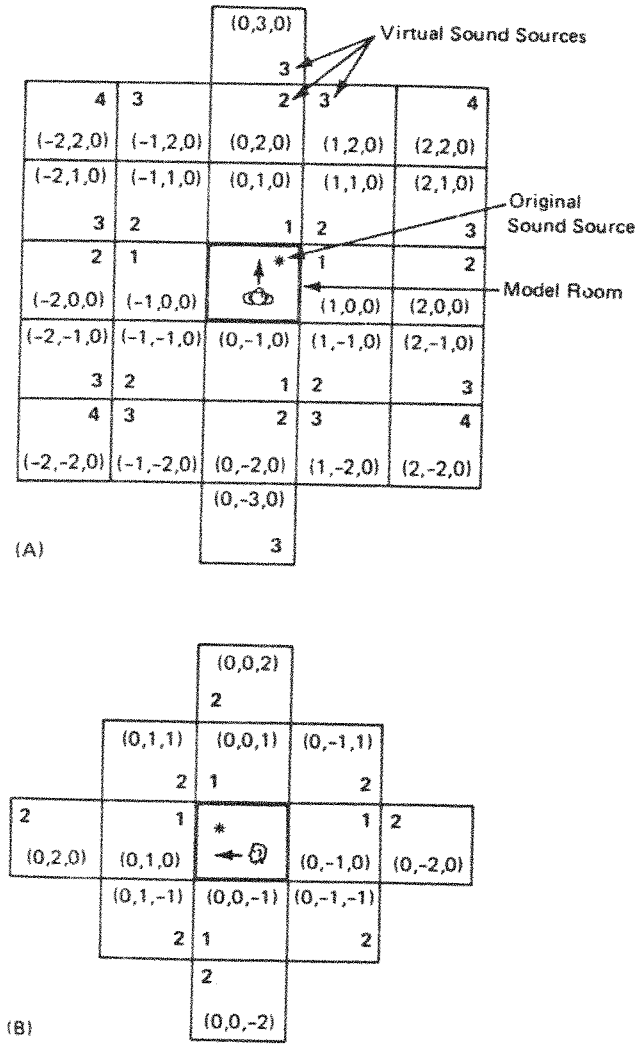


Figure 7.5
Two-dimensional cross section of virtual rooms: (A) horizontal plane and (B) Median plane.

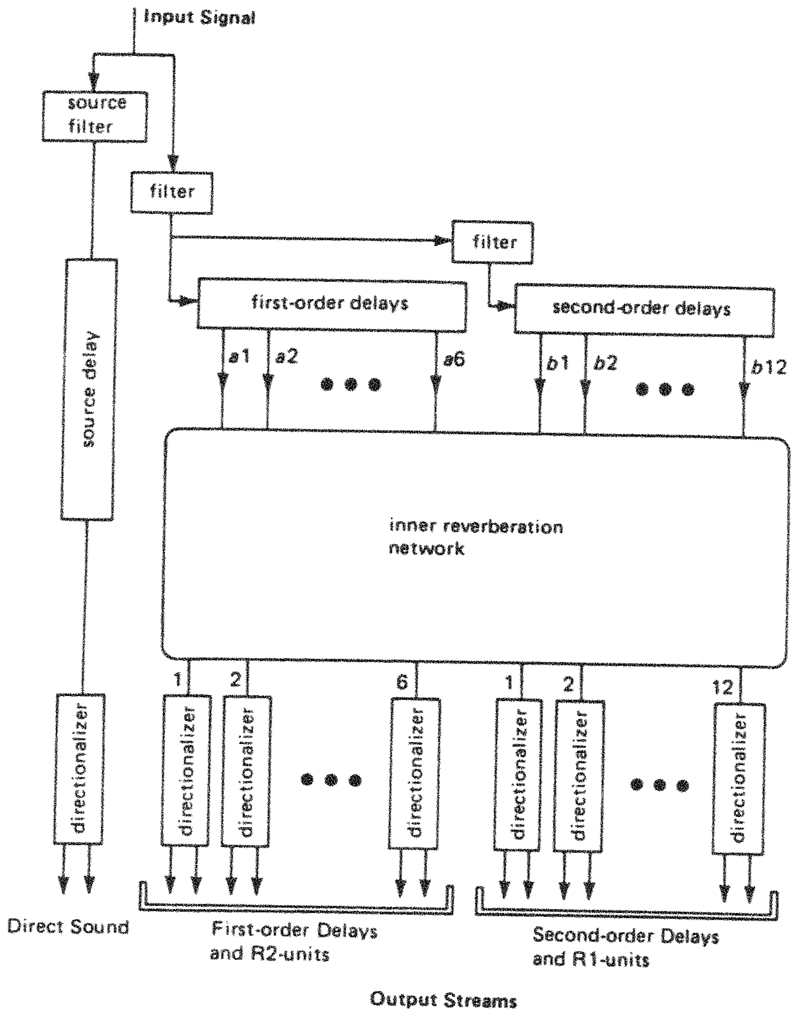


Figure 7.6
Signal-processing network for spatial reverberation as implemented in *space18* program.

model that emanate from virtual rooms behind the six walls of the model room. These first-order virtual sources are contained in the following virtual rooms:

$(1,0,0)\quad(0,1,0)\quad(-1,0,0)\quad(0,-1,0)$

$(0,0,1)\quad(0,0,-1)$

The gain of each reflection is produced by multiplication with the scaling coefficients, $a1$ through $a6$. The signal from each of these delay taps is passed to the inner reverberation network.

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 and will be produced within the inner reverberation network. The remaining twelve second-order delays originate in second-order virtual rooms that extend from the junction of two walls in the model room. Twelve taps from the second delay buffer replicate the time delays for these reflections, and the gain of each reflection is produced by multiplication with the scaling coefficients, $b1$ through $b6$. The signals are passed directly into the inner reverberation network, where they are used to generate reverberation streams that begin with these second-order reflections. These second-order virtual sources are contained in the following virtual rooms:

$(1,0,1)\quad(0,1,1)\quad(-1,0,1)\quad(0,-1,1)$

$(1,1,0)\quad(-1,1,0)\quad(-1,-1,0)\quad(1,-1,0)$

$(1,0,-1)\quad(0,1,-1)\quad(-1,0,-1)\quad(0,-1,-1)$

The exact delay and direction of each reflection is computed from the position of the listener in the model room and the position of the virtual sound source.

Figure 7.7 shows a two-dimensional slice of image rooms for the horizontal plane. Virtual sound sources in the darkly shaded virtual rooms are captured by the initial nonrecirculating delays.

Reverberation Units. There are two types of recirculating delay units within the reverberation subsystem, which will be referred to as the R1-unit and the R2-unit. Both units include a path for the input signal from the delay buffers to be added directly into the output; this path passes the first- and second-order reflections from the nonrecirculating delays into the

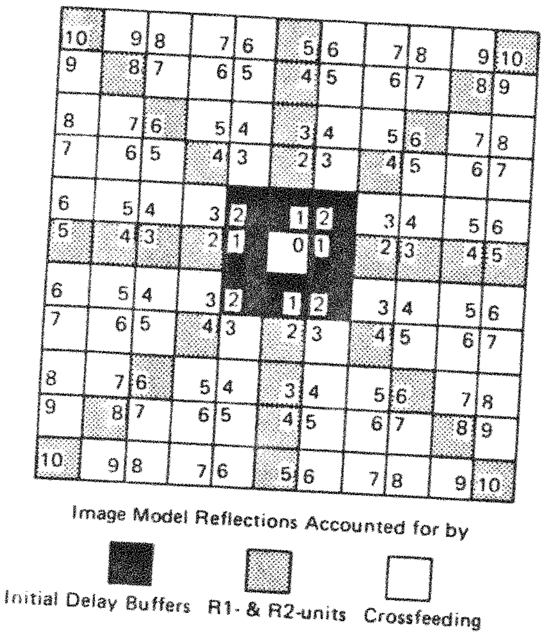


Figure 7.7
Two-dimensional cross section of virtual images rooms. The shading indicates those virtual rooms accounted for by initial delay buffers, R1- and R2-units, and crossfeeding.

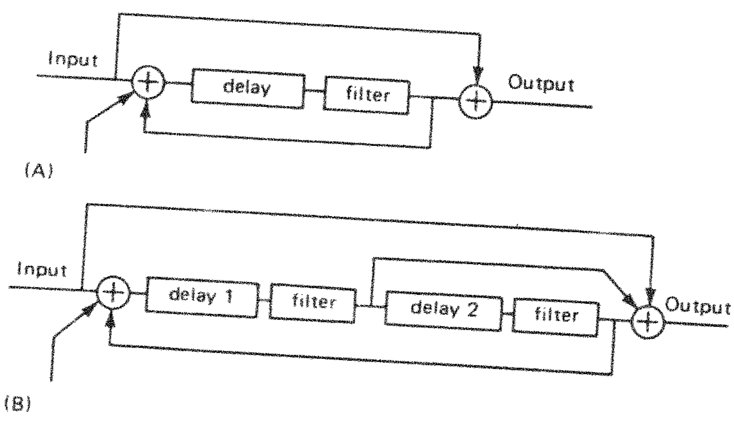


Figure 7.8
Reverberation units: (A) R1-unit and (B) R2-unit.

reverberation streams. There is also an input for signals generated in the crossfeeding process (described below) to be passed into the recirculation. The remainder of the R1-unit is a recursive comb filter similar to that discussed by Schroeder [15]. Schroeder's original version contained a delay buffer and a feedback loop. The amount of feedback was governed by a feedback coefficient in the loop. Moorer [10] implemented a digital realization of Schroeder's suggestion for a RC-section in the feedback loop by incorporating a one-pole, low-pass filter. The response of the filter in the R1- and R2-units is scaled by the attenuation factor for the feedback. Other aspects of the filter are discussed below. The design presented in figure 7.8 places a feedback filter at the end of the delay buffer, but in other respects mimics the Schroeder design. As shown in figure 7.8B, the R2-unit contains a pair of delay buffers with a feedback filters. The actual feedback occurs after the second delay buffer and its feedback filter. The output of the unit is the sum of the outputs of each delay-buffer pair after filtering. This unit produces a pattern of alternating long and short delays that is essential to capturing image model reverberation.

A signal-processing network for spatial reverberation requires a different structure from the Schroeder reverberator, because it must produce parallel streams of reverberation. Schroeder, Moorer, and others use combinations of reverberation units in parallel and series with the final output mixed down to a single reverberation stream. Even in cases where the processing path separates at the end and distinguishes the reverberation streams sent to the individual reproduction channels, the basic combination of reverberation units is in series. The general solution to the problem of producing multiple streams is to place reverberation units in parallel, where each unit produces a unique reverberant stream.

Each of the delay taps from the buffer for first-order reflections is fed into the input of a R2-unit. Each R2-unit is associated with a reverberant stream emanating from a second-order virtual room directly behind the first-order room. For example, a second-order room (2, 0, 0) is directly behind a first-order room (1, 0, 0). The delay lengths in the R2-units are taken from the time of arrival difference of first- and second-order reflections and of second- and third-order reflections, respectively. For the unit associated with room (2, 0, 0), the delay times are given by

$$\text{delay\_1} = T(2,0,0) - T(1,0,0),$$

$$\text{delay\_2} = T(3,0,0) - T(2,0,0),$$

where $T(x, y, z)$ is the predicted time of arrival for a virtual sound source from the virtual room (x, y, z). As shown in figure 7.7, the pattern of reflections emanating from image rooms behind the first-order rooms has a characteristic pattern of alternating short and long delays, especially when the sound source is near a wall. The R2-unit recreates this pattern and thus captures the reverberation for the series of image rooms extending from every wall in the model room.

Each of the delay taps from the buffer for second-order reflections is fed into the input of an R1-unit that is associated with a reverberation stream emanating from a fourth-order virtual room directly behind a second-order room at the junction of two walls. For example, the fourth-order room (2, 2, 0) is directly behind second-order room (1, 1, 0). As shown in figure 7.7, the pattern of reflections emanating from the wall junctions also demonstrates this pattern of short and long delays. However, the long/short pattern is only exaggerated when the sound source is in the corner, and the reflection stream from the wall junction becomes high order twice as fast as the wall reflection stream. Therefore, for the sake of efficiency we chose to implement these delays with the R1-unit, even though it cannot produce a long/short delay pattern. (If greater accuracy is desired, an alternative realization of the inner reverberation network can be created using R2-units in place of the R1-units.) The time delays for the R1-units are taken from the time of arrival difference of second- and fourth-order reflections. For the unit associated with virtual room (1, 1, 0), the delay times are given by

$$\text{delay} = T(2, 2, 0) - T(1, 1, 0).$$

The lightly shaded virtual rooms in figure 7.7 are accounted for by the R1- and R2-units. Together, the eighteen reverberation units produce reverberation streams for eighteen directions in three-dimensional space. Six streams emanate from walls, and twelve streams emanate from the junction of walls. Figure 7.9 shows all the delays calculated for a room 6.2 meters by 12.3 meters by 8 meters.

Crossfeeding of Reverberation Units in Parallel Combination. The elements of the inner reverberation network explained so far replicate all reflections originating in the eighteen lines of virtual rooms extending behind each wall and behind each junction of two walls. Even though this accounts for a large number of reflections, it omits those reflections pre-

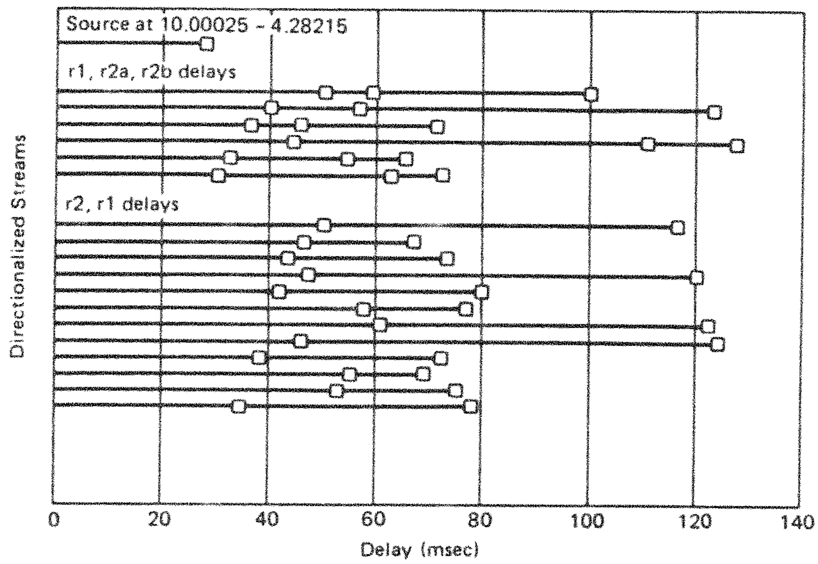


Figure 7.9
Delays calculated for a room 6.2 meters by 12.3 meters by 8 meters. Represented are the source delay, first-order reflection and R2-unit delays, and second-order reflection and R1-unit delays.

dicted by the image model that originate in virtual rooms that lie between those eighteen directions. Close to the source room, there are very few missing rooms. As reflections emanate from farther away in higher-order image rooms, the number of missing rooms greatly outnumbers the rest. Without the "in-between" rooms, the density of reflections does not increase with time.

In order to capture the missing reflections, the output of the R1-units must be fed into the R2-units for spatially adjacent streams. The crossfed signal is added into the initial summation node for each unit. Figure 7.10 represents the crossfeeding process for a single quadrant of a two-dimensional plane. Figure 7.10A shows all of the image rooms up to the fifth order for this region; the source room is to the left. The sequence of reflections emanating from the image room behind the wall, (1, 0, 0), is captured by the R2-unit. The sequence of reflections emanating from rooms behind the two wall junctions, (1, 1, 0), and (1, −1, 0), are produced by R1-units. When the output of each R1-unit is fed immediately into the R2-unit, each reflection is delayed by the R2 delays. The R2 delays are

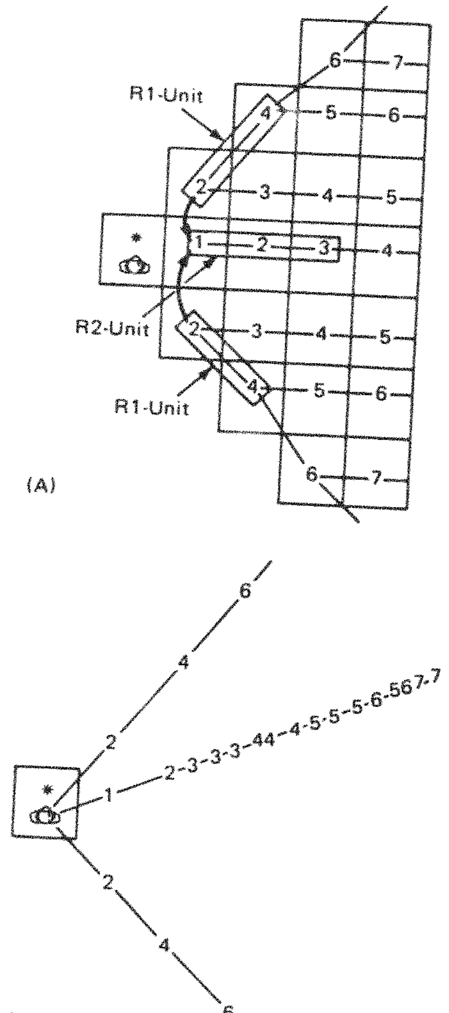


Figure 7.10
Two-dimensional crossfeeding: (A) interaction of R2- and R1-units and (B) the resulting reverberation streams.

approximately equal to the delays between the R1 reflections and those from the next adjoining image rooms to the right. For example,

$T(2,1,0) \simeq T(1,1,0) + \text{R2\_delay 1},$

where R2_delay 1 $= T(2,0,0) - T(1,0,0)$. As the R2-unit recirculates its input, it creates the sequence of reflections whose delay is approximately equal to those emanating from the next set of image rooms between the R1 and R2 rooms. For example,

$T(3,1,0) \simeq T(1,0,0) + \text{R2\_delay 1} + \text{R2\_delay 2},$

where R2_delay 2 $= T(3,0,0) - T(2,0,0)$. As the process continues, the number of reflections is exactly that predicted by the image model. The delay times are entirely accurate for first- and second-order reflections; reflections beyond the second are approximately correct. The output of each R2- and R1-unit can be directionalized toward the leading first- or second-order reflection; the reverberation streams produced are those shown in figure 7.10B.

The crossfeeding illustrated in figure 7.10 is easily extended to all three dimensions. The R2-units are each crossfed from two spatially adjacent R1-units in the horizontal two-dimensional plane and are crossfed themselves into the R1-units that are spatially adjacent in the vertical two-dimensional plane. For example, the R2 stream emanating from the virtual room (1, 0, 0) is spatially adjacent to R1-units associated with the second-order virtual rooms: (1, 0, 1) and (1, 0, −1). Reflections are created for all of the "missing" virtual image rooms from all three dimensions (with the exception of eight third-order rooms). A more intuitive understanding of how the system operates can be gained from figure 7.11, which shows a three-dimensional representation of the network. The *space18* program provides accurate spatial placement for the source and the initial reflection in each reverberation stream, i.e., a total of nineteen sound directions.

7.4 Software for Controlling Spatial Reverberation

To simulate a model listening environment, the user must specify a spatial configuration consisting of 11 parameters: room size (3 dimensions), sound source location (3 coordinates), listener location (3 coordinates), and listener orientation (the direction the listener is facing, specified by azimuth and elevation angles). All of these parameters may be dynamic; the source

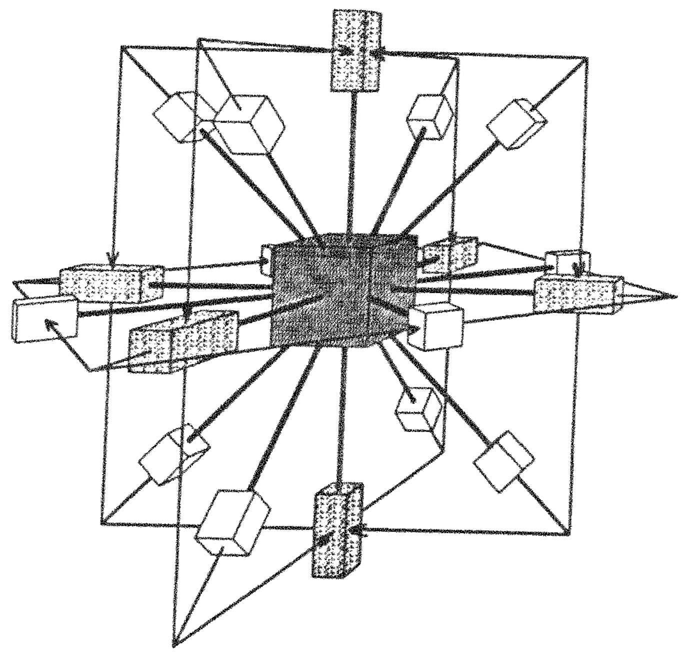


Figure 7.11
Three-dimensional configuration of reverberation units. Crossfeeding signal paths are indicated for the closest R2-unit.

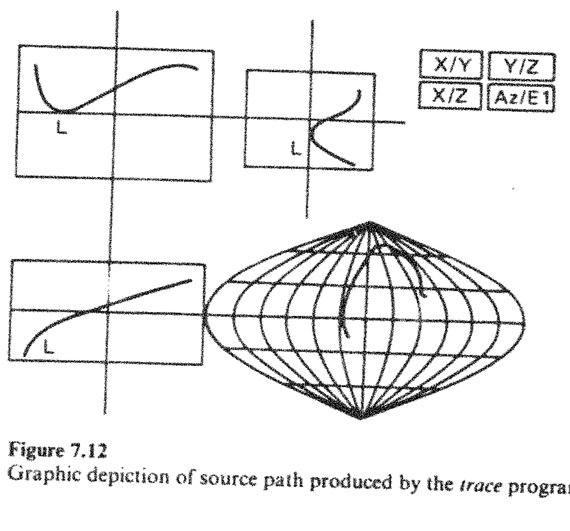


Figure 7.12
Graphic depiction of source path produced by the *trace* program.

and listener may follow an arbitrary path, and room dimensions may also change arbitrarily (the walls of the room follow a "path" as well). The first of the control programs, *framer*, samples the paths for each parameter at regular time intervals determined by a *frame rate*. Its output is a series of *frames* that contain a single value for each of the 11 parameters at a point in time. All subsequent programs, including the reverberator, operate on a frame-by-frame basis.

The user defines paths for all parameters by a list of target values to be reached at particular times, called "anchors," and intermediate values that control curvature, called "attractors." Bezier curves (Rogers and Adams [13]) are used to interpolate a smooth path. The user develops the spatial paths by graphical interaction with the program *trace*. A sample screen is shown in figure 7.12. Once the path has been specified, the program *framer* generates a series of snapshots of the interpolated path at the frame rate.

The first step in modeling a spatial configuration is to determine the spatiotemporal distribution of reflections predicted by the physical model. *Image18*, the second control program in the pipeline, solves the image model (see below) for the spatial configuration and calculates delays and directions for a number of reflections. As described below, the signal-processing network of the reverberator approximately captures high-order reflections with recirculating delays, and so only a small number of reflections must be modeled exactly. The output of *image18* is a series of frames, each of which specifies delays and directions of reflections at a point in time.

7.5 Conclusion

The spatial reverberator is the first reverberation system that recreates the full spatiotemporal sound field of a natural environment. This reverberator is also the first using recirculating delays to create reverberation based on the image model. It is our intention that the physical modeling of acoustic space will provide a point of departure for a more perceptually oriented study of spatial hearing and its relationship to computer music and audio reproduction. A physical model is only useful to the extent that it aids the user in realizing creative intentions. A great deal still needs to be learned about the relationship of spatial sound perception to music perception in general. It is our hope that spatial reverberation will be a tool in this study.

Acknowledgments

This work has been supported by a grant from the System Development Foundation. Special thanks must be given to Charles Smith, without whose commitment to new ideas this project would not have been possible. We also give our thanks to Carl York, whose continuing interest and support has meant a great deal to us.

References

[1] Allen, J. B., and Berkeley, D. A. (1979). Image method for efficiently simulating small room acoustics. *JASA*, 65(4), 943–950.

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

[3] Blauert, J. (1969). Sound localization in the median plane. *Acustica*, 22, 957–962.

[4] Blauert, J. (1981). Lateralization of jittered tones. *Journal of the Acoustical Society of America*, 70, 694–698.

[5] Blauert, J. (1982). *Spatial Hearing*, trans. John S. Allen. MIT Press (Cambridge, MA). Originally entitled *Raumliches Hören*, S. Hirzel Verlag (Stuttgart), 1974.

[6] Butler, R. A., and 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.

[7] Chowning, J. M. (1971). The simulation of moving sound sources. *Journal of the Audio Engineering Society*, 19, 2–6.

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

[9] Moore, F. Richard (1983). A general model for spatial processing of sounds. *Computer Music Journal*, 7(3), 6–15.

[10] Moorer, James A. (1979). About this reverberation business. *Computer Music Journal*, 3(2), 13–28.

[11] Musicant, A. D., and Butler, R. A. (1980). Monaural localization: an analysis of practice effects. *Perception and Psychophysics*, 28, 236–240.

[12] Nuetzel, J. M., and 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.

[13] Rogers, D. F., and Adams, J. A. (1976). *Mathematical Elements for Computer Graphics*, 139–144.

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

[15] Schroeder, M. R. (1962). Natural-sounding artificial reverberation. *Journal of the Audio Engineering Society*, 10, 219–223.

[16] Schroeder, M. R. (1970). Digital simultion of sound transmission in reverberant spaces. *JASA*, 47(2), 424–431.

[17] Shaw, E. A. G., and 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, 240–249.

[18] Sheeline, C. W. (1982). An investigation of the effects of direct and reverberant signal interaction on auditory distance perception. PhD dissertation, Department of Hearing and Speech Sciences, Stanford.

[19] Smith, J. O. (1984). An all-pass approach to digital phasing and flanging. *Proc. ICMC*, 1984, 103–109.

[20] Wright, D., Hebrank, J. H., and Wilson, B. (1974). Pinna reflections as cues for localization. *Journal of the Acoustical Society of America*, 56, 957–962.