1. INTRODUCTION

The "spatial reverberation" process (Kendall & Morris, 1981; patent pending) attempts to duplicate the full range of spatial listening cues present in natural acoustical environments. The cues are largely of two types. First, there are cues for directional hearing based on interaural time difference, interaural intensity difference, and the spectral shaping of the upper tonal, loud and quiet cues (lower ear). Second, there are cues for distance and room acoustics based on accurate modeling of experimential reverberant sound. It is the combination of directional cues with environmental modeling that enables "spatial reverberation" to support the sensation of spatiotemporal pattern of sound in natural environments.

In contrast is the original Schauwer (1966) reverberator, which has a single control over the reverberation time. "Spatial reverberators" require a detailed specification of the acoustical characteristics of the simulated environment. It is also necessary to specify the dynamic relationship of sound objects to the listener (in a requirement stated by the spatial processing systems developed by Cowles (1971) and Holm (1983)). The process of specification is aided by having an underlying model of the physical setting. The user can dynamically manipulate such parameters as room dimensions, sound source location, and listener location, while other parameters such as the absorption coefficients of the walls, ceiling, and floor. Although control of spatial reverberation will eventually be oriented toward psychological specification of the subjective dimensions of spatial sound images, physical models provide a starting point for the exploration of creative potentials and for research into the underlying organization of the psychological dimensions.

This paper provides an overview of the spatial processing software developed at Northwestern Computer Music (NCSM) during the past two years. We begin by discussing "control software" for specifying reverberation characteristics and then proceed to describe the signal processing inside the reverberator itself.

2. SOFTWARE FOR CONTROLLING REVERBERATION

There is a natural necessity by which the data necessary to control "spatial reverberation" must be created. This sequence starts with the high-level description of the shape of the environment and the position of the sound source and listener within that environment and ends with a comprehensive list of their delays, gain factors, filter coefficients, etc. Low-level information is converted into low-level information in several discrete steps performed by individual programs in series. Fig. 1 diagrams the flow of data through these programs. Although the manipulation of this data would be more efficient if contained in a single program, the software has been split into separate phases; since development has slowed and the production of music has taken over, these programs will be consolidated and made more efficient. Our modular approach has enabled us to expand and develop programs quickly because each program is relatively small and performs a clearly circumscribed set of spatial tasks.

![Figure 1. The flow of data through control software.](image-url)
2.1 Specifying a Spatial Configuration

To simulate an actual scanning environment, the user must specify a spatial configuration consisting of 11 parameters: room size (5 dimensions, ground-source level (3 coordinates), interior location (3 coordinates), and boundary orientation (the direction the former is facing, specified by azimuth and elevation angles). All of these parameters may be dynamic: the source and target may follow arbitrary paths, and room dimensions may also change arbitrarily (the walls of the room follow an "path" as well). The firm of the control program, frames, sample the path for each parameter at regular time intervals determined by a frame rate. Its output is a series of frames, each containing a single sample 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 the path for all parameters by a list of target values to be reached at particular times, called "targets" or "intermediate values" which control curves, called "targets." Several scenes are then inserted in the program, creating a series of snapshots of the interpolated path at the frame rate. No further specifications are required from the user. The remaining control programs calculate reverberation parameters to simulate the desired spatial configuration.

Figure 2. Graphic depiction of source path produced by the scene program.

3.2 Modeling Spatial-Temporal Distribution

The first step in modeling a spatial configuration is to determine the spatial-temporal distribution of reflections predicted by the physical model. As described below, the signal processing network of the reverberator is essentially a high-order reflection with reverberating delays, and so it is a small number of parameters that are meaningful. The output of equation X is a series of frames, each of which specifies delays and directions of reflections at a point in time.

3.3 Modeling Energy Losses

The decay and direction information from equation Y describes the spatial-temporal distribution of reflections; the first step is to model the energy losses associated with each reflection. As described below, each delay unit within the reverberator is in series with an attenuation factor (gain) and a 2-tone filter. The attenuation factor simulates a frequency-independent energy loss, while the filter simulates a frequency-dependent energy loss. The third and final control program in the pipeline, equation Z, provides the user with control over the simulation of attenuation factors and filter coefficients.

Attenuation of early reflections is weighted by an arbitrary constant. Attenuation factors for reverberating delays are calculated from this constant or from user-specified reverberation rates. The user-weighting constant is appropriate for small room acoustics, while for reverberation in an appropriate for large room acoustics. Low-order filters model the frequency-dependent energy absorption by the air and the walls of the room. The user only needs to specify the extent and the frequency-dependency of air and wall absorption using standard absorption coefficients.

3.4 Reverberation Signal-Processing

3.2 Recreating Delayed and the Image Model

The technique of recreating delayed and high-order timing and digital audio production is one of the major contributions of Manfred Schroeder's pioneering work in 1972. The Schroeder reverberator produces small units of reverberating delays combined in a predominantly serial network. Schroeder explains the reason for choosing the serial configuration:

"Previous investigators have suggested multiple feedback to produce a higher echo density. However, multiple feedback has severe stability problems. Also, it leads to numerous frequency resonances and non-erential decay characteristics. A much easier solution to this problem would be to add a suitable unit which could connect in series any desired number of times. In this manner, each unit would effectively realize the number of echoes produced by the preceding unit."

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 together produce a number of echoes akin to that measured in real reverberation. Schroeder also suggests that the feedback coefficients within the comb filters be made frequency-dependent.

"If it is desired to track the reverberation time as a function of frequency, ... at a single SC-section in each feedback path will suffice. In this manner further results can be added to the spectral reverberation by using the reverberation time longer for lower frequenc."
In 1970, Schroeder expanded his reverberation network to include an initial delay buffer which replicates the sound of the early reverberation pattern typical of concert halls. This model, including both the frequency-dependent reverberation delays and the simulation of early reflections, was refined by Moore (1971) who produced the time-weening reverberation network of this type. A significant aspect in the design of such networks is achieved by Christian (1970), who combined the basic model of reverberation delays with time-weening controls to simulate moving sound sources.

This and other contemporary suggestions in reverberation attempts to replicate the global reverberation typical of large reverberation 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 move a room, simulating the reverberation in the characteristics of a particular room is largely a matter of perspective. Even Christian'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.

The "image model" of reverberation (Diles and Deller, 1973) provides a method for predicting both the spatial and the temporal patterns of reflected sound from the room dimensions and the positions of the acoustic sources and the listener within that room. For this model, each ray of reflected sound is visualized as originating from a "virtual sound source" outside of the actual physical room. Each virtual sound source is positioned within a "virtual room" that replicates the physical room or in a mirror image of it. We use three-dimensional coordinates (x,y,z) to specify individual virtual sound sources or virtual rooms (Fig. 2). The pattern of reflected sound in the physical room can be visualized as the compound sound reaching the listener from all virtual sources. A computer measuring the position of the listener in the physical room and these virtual sources will predict the pattern from which the reflected sound emanates and the distance that the sound must travel before reaching the listener. Even though real sound waves have irregularities that cause reflections to be easily difficult, the image model provides a good approximation of the direction and timing of the more important reflections.

Until now, the image model was computationally too expensive to solve for sound production alone. If could only be implemented by convolution with an early stimulated sound impulse response. It has been generally believed that the image model could not be explored with computer delays. For example, Kooper (1978) states:

"The only problem with these system of reverbberation is that they are not comntended in the same way as reverbberation networks for real rooms. This is because the reflection is in a room, not a square one, do not typically come in regular sequences separated by equal amounts of time."

![Figure 2. Three-dimensional representation of a virtual scene.](image_url)
3.2 Signal-Processing Overview

Although the Kendall-Kuipers network was conceived in part of "spatial reverberation," the design is applicable to reverberation simulation in general and can be applied in many different ways. A basic signal-processing network for spatial reverberation requires two subsystems (Fig. 4). One is a reverberation subsystem that takes an input signal and produces multiple outputs, each of which is a different "reverberation stream." The other is a directional analysis subsystem that takes an input signal, superimposes directional cues, and produces multiple outputs. These directional cues might involve pan- oiling, the output from a network speaker, but more likely involve superimposing horizontal cues. Our discussion will focus on the reverberation subsystem.

![Diagram of signal-processing network for spatial reverberation.](image)

The input signal to the spatial reverberator is sent to the first directionalizer subsystem and to the reverberation subsystem. The first directionalizer subsystem determines the linear direction of the unmodified input signal. The reverberation subsystem produces the input signal and produces a number of "reverberation streams." Each of these reverberation streams captures the temporal pattern of reflected sound coming to the listener from a specific region of the room. These streams are then sent to directionalizer subsystems which determine the linear direction of each reverberation stream. The output signals from these directionalizer subsystems are mixed together to create a composite of the input signal and the reverberation streams which create a three-dimensional sound field. The ensemble of reverberation streams captures the spatial-temporal patterns of reflected sound predicted by the room model. The number of separate audio channels in the composite output is determined by the number of channels available for the sound reproduction, but should be at least two in order to preserve different signals to the listener's left and right ears.

3.3 Delays for the Source and First- and Second-Order Reflections

This basic signal-processing network can be tailored in a number of ways. Fig. 5 illustrates the version of the system implemented for the spatial program currently in use at Northeastern University. The input signal to \( \text{spatial} \) is fed into three different non-redundant delay buffers. Digital signal interpolation must be performed on all buffers when the delay lines change (Stein 1981). The low buffer performs delays for the source itself. Its input is mixed and filtered to capture intensity and spectral changes due to distance \\( \text{or} \) to a discriminator. The source signal enters the \\( \text{spatial} \) network through three delay buffers with multiple taps. Each buffer is preceded by a filter which captures the spectral changes due to air \\( \text{and} \) to a discriminator. The output of these filters produces the delays for the six first-order reflections predicted by the room model without echoes from virtual sources behind the six walls of the student room. Three sinusoidal virtual sources are contained in the following virtual room:

\[
\begin{align*}
(1, 0, -1) & \quad (0, 1, 0) \\
(1, 1, 0) & \quad (0, 0, 0) \\
(0, 1, 1) & \quad (0, 0, 1)
\end{align*}
\]

The gain of each reflection is produced by multiplication with the scaling coefficients, \( g(s) \). The signal from each of these delay taps is passed to the \\( \text{spatial} \) reverberation network.

![Diagram of signal-processing network for spatial reverberation.](image)
The image model predicts a total of eighteen second-order reductions. Set of fans 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 eighteen second-order delays originate in second-order virtual rooms that stem from the junction of two walls in the model room. Twelve taps from the second delay buffer replace the time-delays by class reflections and the path of each reflection is retrieved by multiplication with the scaling coefficients, 0.5 and 1.0. The signals are passed directly into the inner reverberation network where they are used to generate reverberation streams that begin with class second-order reflections. Three second-order virtual sources are contained in the following virtual rooms:

- (1, 0, 1)  
- (0, 1, 1)  
- (0, 0, 1)  
- (0, 1, 0)  
- (1, 0, 0)  
- (1, 1, 0)

The start 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.

Fig. 6 shows a two-dimensional slice of image rooms for the longitudinal plates. Virtual sound sources in the darkly shaded virtual rooms are captured by the initial non-reverberating delays.

3.4 Units Within the Inner-Reverberation Network

There are two types of reverberating delay units within the reverberation subsystems which will be referred to as the "R1-unit" and the "R2-unit." Both units include a path for input signals to be added directly into the output; this path passes the first- and second-order reflections from the non-reverberating delays into the reverberation stream. The remainder of the R1-unit's a recursive comb filter similar to that discovered by Schröder (1962). Schröder's original version contained a delay buffer and a feedback loop. The amount of feedback was governed by a feedback coefficient in the loop. However, Sherr (1974) implemented a digital realization of Schröder's suggestion for a R2-section in the feedback loop by incorporating a two-pole low-pass filter. The response of the filter in the R1- and R2-units is varied by the attenuation factor for each feedback. These factors for the filter in denoted below. The design presented in Fig. 7a places a feedback filter at the end of the delay buffer, but in other respects follows the Schröder design. As shown in Fig. 7b, the R2-unit contains a pair of delay buffers into a feedback loop. The actual feedback occurs after the second delay buffer and its feedback filter. The output of the tail in the rear of the outputs of each delay buffer pair after filtering. This unit produces a pattern of alternating long and short delays which is essential in capturing image model reverberation.

(a)  

(b)  

Figure 7. Reverberation units: (a) R1-unit; (b) R2-unit.

A signal-processing network for spatial reverberation requires a different structure than the Schröder reverberation, because it must produce parallel streams of reverberated, Schröder, blower and other sound combinations of reverberation units is parallel and occurs with the final output mixed down to a single reverberation stream. Even in cases where the processing path network in the end and distinguishes the reverberation stream sent to the left- and right-reproduction channels. The basic combination of reverberation units is shown. The general solution is the problem of producing multiple reverberation units in parallel where each unit produces a unique reverberation stream.

Each of the delay taps from the buffer for second-order reflections is fed into the input of a R2-unit. Each R2-unit is associated with a reverberation stream consisting from a second-order virtual room directly behind the first-order room. For example, second-order room (1, 0, 1) is directly behind first-order room (1, 0, 0). The delay samples in the R2-unit are taken from the timearrivaldifference of first- and second-order reflections and of second- and third-order reflections respectively.

For the unit associated with room (1, 0, 0), the delay times are given by:
The delay time from the buffer for second-order reflection is fed into the input of a first-order unit which 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 four-order room (3, 2, 0, 0) is directly behind second-order room (1, 1, 0). As shown in Fig. 2, the pattern of reflections emanating from the wall junction also determines that pattern of short and long delays. However, the downward 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 clarity we choose to implement these delays with the S1-unit, even though it is possible to produce a long-dead delay pattern. (If greater accuracy is desired, an alternative realization of the inter-reverberation network can be created using S2-units in place of the S1-units.) The time delays for the S1-units are taken from the time of travel distance of second- and fourth-order reflections. For the unit associated with virtual room (1, 1, 0), the delay times are given by:

\[ \text{delay} = [T_1, 2, 0] \cdot (T_1, 1, 0) \]

The lightly shaded virtual rooms in Fig. 6 are all second-order S2-units and S1-units. Together, the 18 reverberation units produce reverberation streams for what directions are three-dimensional space. Six streams emanate from walls, and twelve emanate from the junction of walls. Fig. 4 shows all the delays calculated for a room 6.2 by 12.3 by 6.0 meters.

2.5 Crossfeeding of Reverberation Units in Parallel Combination

The element of the inner reverberation network explained so far replicates all reflections originating in the 18 units in a reverberation stream extending backward and forward behind each junction of two walls. Even though this accounts for a large number of reflections, it omits those reflections predicted by the model image that originate in second-order rooms that lie between those 18 directions. Close to the source room, there are very few missing sounds. As reflections emanate from other area in higher-order image rooms, the number of missing sounds greatly outnumber the rest. Without the "backtracking" room, the density of reverberation does not increase with time.

ICMC 86 Proceedings

290

Figure 8. Delays calculated for a room 6.2 by 12.3 by 8 meters. Represented are the source delay, first-order reflection and S2-unit delays, and second-order reflection and S1-unit delays.

In order to capture the missing reflections, the output of the 31 units must be fed into the 32-unit for spatially adjacent streams. The crossed signal is added into the initial connection made for each unit. Fig. 4 represents the cross feeding process for a single quadrant of a two-dimensional cross section. Fig. 9a shows all of the image rooms beyond the 5th order for this region; the source room is to the left. The sequence of reflections emanating from behind the two wall junctions are produced by S2-units. The sequence of reflections emanating from the wall is captured by the S1-units. When the output of each S1-unit is fed into the 31st unit, each reflection is delayed by the S1 delay. The second delays are approximately equal to the delays between the S1 reflections and those from the next adjacent image rooms to the right. For example:

\[ T_1 = [1, 1, 0] + T_2, 2, 0 \]

where \( T_2 \) is the delay for the second-order reflection between the two junctions.

The S1 units receive their input, they create the sequence of reflections whose delay is approximately equal to those emanating from the next set of image rooms between the 41 and second-order rooms. For example:

\[ T_1 = [1, 1, 0] + T_2, 2, 0 + T_3, 3, 0 \]

where \( T_3 \) is the delay for the second-order reflection between the two junctions.

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 paralled room are approximately correct. The output of each S2 and S3-unit is directionalized toward the leading first- or second-order reflections; the reverberation sound produced are those shown in Fig. 9b. Although the exact spatial position of the "between rooms" is lost, relative spatial position is captured. However, the process is not spatially inverted for the source and the initial reflection in each reverberation stream, i.e., a total of 19 sound directions.
The confounding illustrated in Fig. 2 is easily extended to all three dimensions. The R2-units are each derived from their four spatially adjacent R1-units. For example, the R1 stream emanating from the virtual zone \((1, 0, 0)\) is spatially adjacent to the reverberation stream implemented with R2-units. These R2-units are associated with the second-order virtual zones \((1, 1, 1), (0, 1, 0), (1, 0, 1), (0, 1, 0)\) and \((1, 0, 1)\). Each R2-unit is fed into the two R3-units associated with the adjoining walls. Reflections are required for all of the "missing" virtual image rooms from all three dimensions. A more intuitive understanding of how the system operates can be gained from Fig. 10 which shows a three-dimensional representation of the network. Thus, the evolution in temporal density of reflections is fast predicted by the image model (see Fig. 11). Despite the tremendous increase in reflection density, there is no possibility of the confounding network becoming unstable because there is no feedback.

3.6 Air and Wall Absorption Filters

The original Schlauder reverberator (1923) implements feedback with a reverberation delay box. In order for the reverberation to be stable, the feedback coefficient must be less than 1, i.e., the reverberating signal must lose energy over time. In real rooms, the loss of sound energy in the natural reverberation is frequency-dependent. The air absorbs high frequencies at the sound travels the walls almost high frequencies and low frequencies that escape without being reflected. Moore (1970) uses a one-step filter to separate the general high-frequency loss due to air absorption. This filter continuously shifts the signal being reverberated in a frequency-dependent manner and performs reverberation that becomes progressively more low-pass through time.
4. CONCLUSION

The spatial reverberator is the first reverberation system which recreates the multidimensional sound field of a natural environment. This reverberator is also the first using extraneous delays to create reverberation based on the image model. Having a model simplifies the user's task of specifying intentions, but it also requires an extensive amount of software support in order to perform and manage a large amount of data about the details of the structural space. 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 the perception of sound and audio reproduction. A physical model is only useful if the extent that it aids the user in realizing creative intentions. A great deal still needs to be learned about the relationship of visual versus perception to music perception in general. 's is our hope that spatial reverberation will be a tool of this study.

ACKNOWLEDGMENTS

The work has been supported by a grant from the Syrinx Development Foundation. Special thanks must be given to Charlie Smith without whose encouragement to new ideas and this project would not have been possible. We also give our thanks to Clive York whose continuing interest and support have meant a great deal to us.

REFERENCES


