IMAGE MODEL REVERBERATION FROM RECIRCULATING DELAYS

2408 (C-17)

Gary Kendall
William Martens
Daniel Freed
Derek Ludwig
Richard Karstens
Northwestern University
School of Music
Evanston, Illinois

Presented at the 81st Convention 1986 November 12-16 Los Angeles, California

AN AUDIO ENGINEERING SOCIETY PREPRINT

Image Model Reverberation from Recirculating Delays

Gary Kendall
William Martens
Daniel Freed
Derek Ludwig
Richard Karstens

Computer Music Studio
School of Music
Northwestern University
Evanston, IL 60201 USA
(312) 491-3895

ABSTRACT

The technique of recirculating delays has served as the basis for reverberation simulation in digital audio production ever since the publication of Manfred Schroeder's pioneering work in 1962. The Schroeder reverberator prescribes small units of recirculating delays in a configuration that is predominantly serial. Extensions to this method by Chowning (1970) and Moorer (1979) have achieved substantial improvements in realism and naturalness, while maintaining the basic structure of the Schroeder model. 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.

The "image model" of reverberation (Allen and Berkeley, 1979) 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 the room. 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. In the process of creating the "spatial reverberation" technique, Kendall and Martens (1984) devised a reverberation network with recirculating delays that captures the spatial and temporal distribution of reflected sound predicted by the image

model. The network consists of recirculating delay units that are connected entirely in parallel with non-recirculating cross feeding. The "reverberation stream" produced by each unit captures the temporal pattern of reflected sound from a specific region of the room. The ensemble of reverberation streams produced by the entire network captures the spatial and temporal pattern of reflected sound predicted by the image model. The output of the reverberator may be mixed to mono, lateralized to stereo, or spatialized into three-dimensions as "spatial reverberation." Experience using the reverberator have shown that it is capable of achieving a high degree of realism with both large and small simulated rooms as well as creating vivid illusions of source distance.

1. INTRODUCTION

The "spatial reverberation" process (Kendall & Martens, 1984; patent pending) attempts to simulate the full range of spatial hearing 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 torso, head and pinna (outer ear). Second, there are cues for distance and room acoustics based on accurate modeling of environmental reflected sound. It is the combination of directional cues with environmental modeling that enables "spatial reverberation" to capture the complete spatiotemporal pattern of sound in natural environments.

2. REVERBERATION SIGNAL-PROCESSING

2.1 Recirculating Delays and the Image Model

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. The Schroeder reverberator prescribes small units of recirculating 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 non-flat frequency responses and non-exponential decay characteristics. A much easier solution to the echo density problem would be at hand if one had a basic reverberating unit which one could connect *in series* any desired number of times. In this manner, each unit would effectively *multiply* the number of echoes produced by the preceding units."

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. Schroeder (1962) also suggests that the feedback coefficients within the comb filters be made frequency-dependent.

> "If it is desired to make the reverberation time a function of frequency, . . . a simple RC-section in each feedback loop will suffice. In this manner further realism can be added to the artificial reverberation by making the reverberation time larger for low frequencies."

In 1970, Schroeder expanded his reverberation network to include an initial delay buffer which replicates the kind of the early reflection pattern typical of concert halls. This model, including both the frequency-dependent recirculating delays and the simulation of early reflections, was refined by Moorer (1979) who produced the best-sounding reverberation network of this type. A significant advance in the goals of such networks was achieved by Chowning (1970), 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.

The "image model" of reverberation (Allen and Berkeley, 1979) 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 (Fig. 1). 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. For example, Moorer (1979) states:

> "The only problem with these systems of recirculating delays is that they can never correspond to the room reflection pattern in a real room. This is because the reflections in a room, even a square one, do not typically come in regular sequences separated by equal amounts of time."

In the process of creating the "spatial reverberation" technique, Kendall and Martens (1984) devised a reverberation network with recirculating delays that captures not only the temporal, but the spatial distribution of reflected sound predicted by the image model. The creation of this network was motivated by a different set of goals than those of Schroeder or Chowning. They state:

> "Our particular experience with reflected sound and pinna cues led us to the design of a reverberation system with different assumptions and goals . . .. We have concluded 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 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."

2.2 Signal-Processing Overview

Although the Kendall-Martens network was conceived as part of "spatial reverberation", the design is applicable to reverberation simulation in general and can be realized in many different ways. A basic signal-processing network for spatial reverberation requires two subsystems (Fig. 2). 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 directionalizer subsystem that takes an input signal, superimposes directional cues and produces multiple outputs. These directional cues might involve panning the output between speakers, but more likely involve using synthesized locational cues. Our discussion will focus on the reverberation subsystem.

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 illusory direction of the unmodified input signal. The reverberator subsystem processes 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 illusory direction of each reverberation stream. The output signals from each directionalizer subsystem 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 spatiotemporal pattern of reflected sound predicted by the image 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 present different signals to the listener's left and right ears.

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

This basic signal-processing network can be realized in a number of ways. Fig. 3 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 non-recirculating delay buffers. Digital signal interpolation must be performed on all buffers when the delay times change (Smith 1984). 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 which 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 model which 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) (0, 1, 0) (-1, 0, 0) (0,-1, 0)
(0, 0, 1) (0, 0,-1)

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.

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, b[1-12]. 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) (0, 1, 1) (-1, 0, 1) (0,-1, 1)
(1, 1, 0) (-1, 1, 0) (-1,-1, 0) (1,-1, 0)
(1, 0,-1) (0, 1,-1) (-1, 0,-1) (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.

Fig. 4 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 non-recirculating delays.

2.4 Units Within the Inner-Reverberation Network

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 non-recirculating delays into the reverberation streams. There is also an input for signals generated in the cross-feeding 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 (1962). 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 (1979) 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 is discussed below. The design presented in Fig. 5a places a feedback filter at the end of the delay buffer, but in other respects mimics the Schroeder design. As shown in Fig. 5b, 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 which is essential to capturing image model reverberation.

A signal-processing network for spatial reverberation requires a different structure than 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, second-order room (2, 0, 0) is directly behind 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 Fig. 4, 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 a R1-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 fourth-order room (2, 2, 0) is directly behind second-order room (1, 1, 0). As shown in Fig. 4, 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 forth-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 Fig. 4 are accounted for by the R1- and R2-units. Together, the 18 reverberation units produce reverberation streams for 18 directions in three-dimensional space. Six streams emanate from walls, and twelve streams emanate from the junction of walls.

2.5 Crossfeeding of Reverberation Units In Parallel Combination

The elements of the inner reverberation network explained so far replicate all reflections originating in the 18 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 predicted by the image model that originate in virtual rooms that lie between those 18 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. Fig. 6 represents the cross feeding process for a single quadrant of a two-dimensional plane. Fig. 6a 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 approximately equal to the delays between the R1 reflections and those from the next adjoining image rooms to the right. For example,

$$T(2, 1, 0) \simeq T(1, 1, 0) + \text{R2\_delay1}$$

where $\text{R2\_delay1} = T(2, 0, 0) - T(1, 0, 0)$. 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,

$$T(3, 1, 0) \simeq T(1, 0, 0) + \text{R2\_delay1} + \text{R2\_delay2}$$

where $\text{R2\_delay2} = T(3, 0, 0) - T(2, 0, 0)$. 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 Fig. 6b. Although the exact spatial position of the "in-between rooms" is lost, relative spatial position is captured. *Space18* provides accurate spatial placement for the source and the initial reflection in each reverberation stream, ie., a total of 19 sound directions.

The crossfeeding illustrated in Fig. 6 is easily extended to all three dimensions. The R2-units that are each crossfed from two spatially adjacent R1-units in the horizontal two-dimensional plane (Fig. 6) 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 Fig. 7 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, ie., a total of 19 sound directions.

3. CONCLUSION

The spatial reverberator is the first reverberation system which recreates the spatiotemporal sound field of a natural environment. This reverberator is also the first using recirculating 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 software support in order to manage the data about the the acoustical 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 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 and 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

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

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

Kendall, Gary S. and Martens, William L. (1984) Simulating the Cues of Spatial Hearing in Natural Environments. Proc. of International Computer Music Conference, 1984, 111-125.

Moore, F. Richard (1983) A General Model for Spatial Processing of Sounds. Computer Music Journal, 7 (3), 6-15.

Moorer, James A. (1979) About This Reverberation Business. Computer Music Journal, 3 (2), 13-28.

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

Schroeder, M. R. (1970) Digital Simulation of Sound Transmission in Reverberant Spaces. JASA, 47(2), 424-431.

Smith, J. O. (1984) An All-pass Approach to Digital Phasing and Flanging. Proc. ICMC, 1984, 103-109.

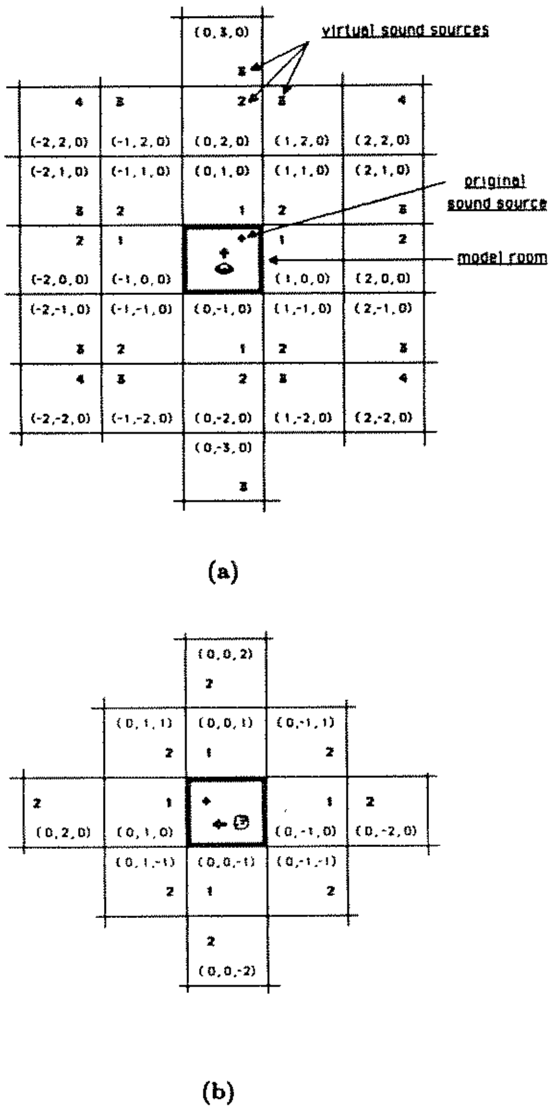


Figure 1. Two-dimensional cross-sections of virtual image rooms. a) horizontal plane, b) median plane.

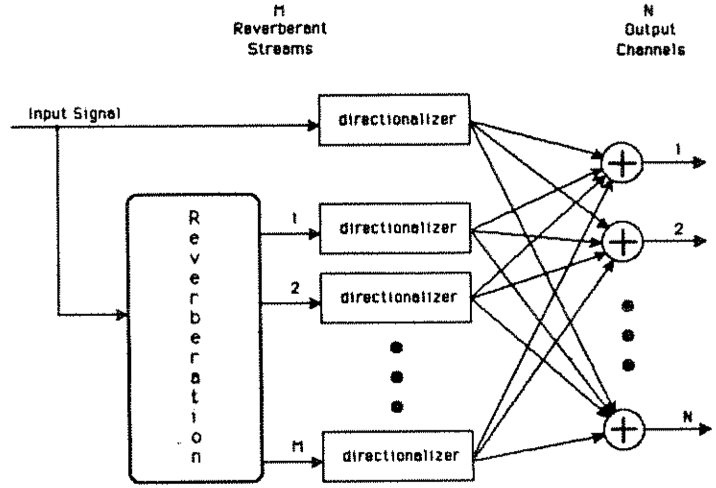


Figure 2. Basic signal-processing network for spatial reverberation.

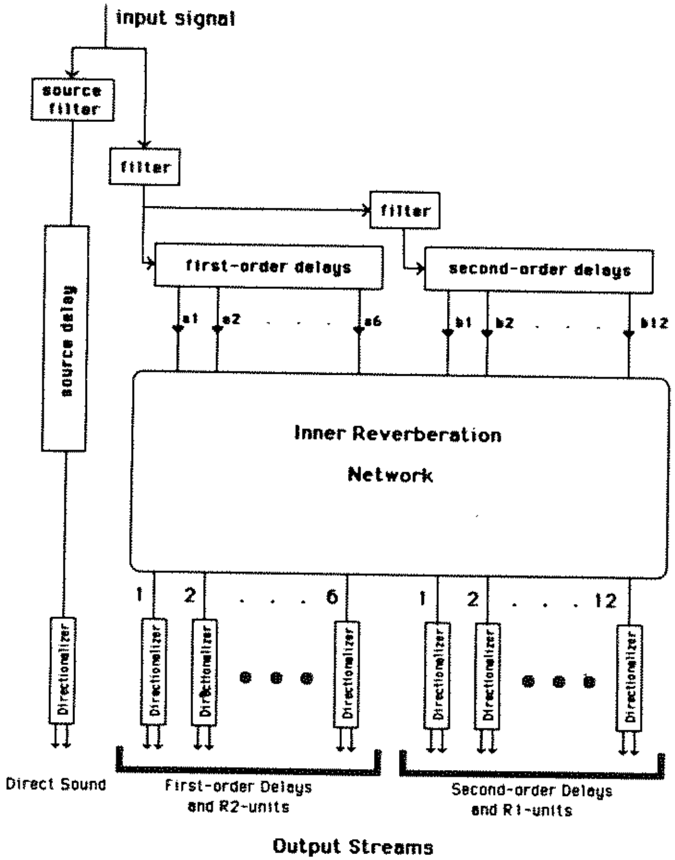


Figure 3. Signal-processing network for spatial reverberation as implemented in *space18* program.

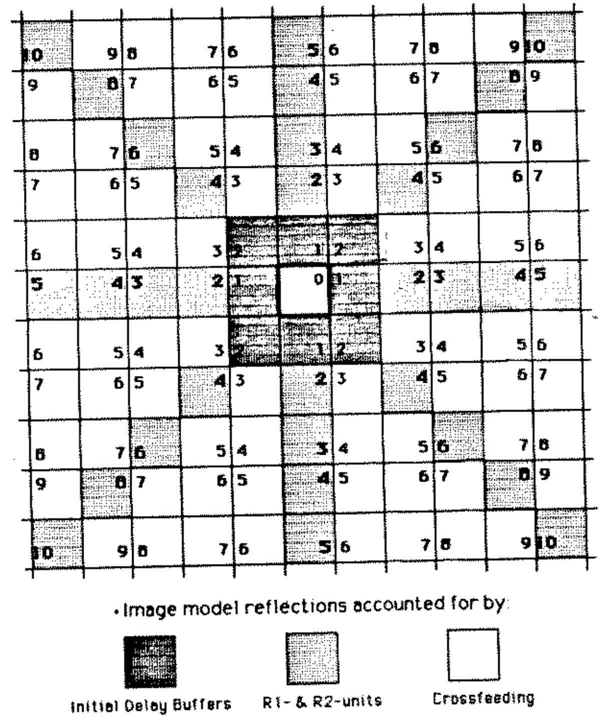
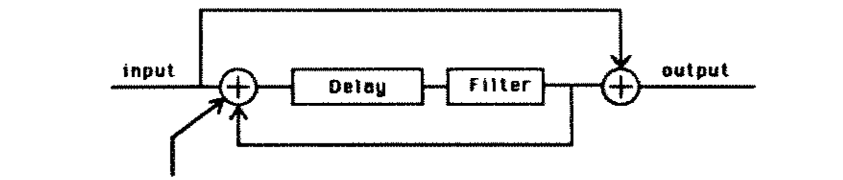
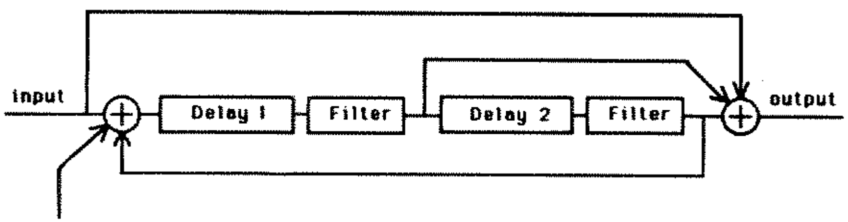


Figure 4. Two-dimensional cross-section of virtual image rooms. The shading indicates those virtual rooms accounted for by initial delay buffers, R1- and R2-units, and crossfeeding.

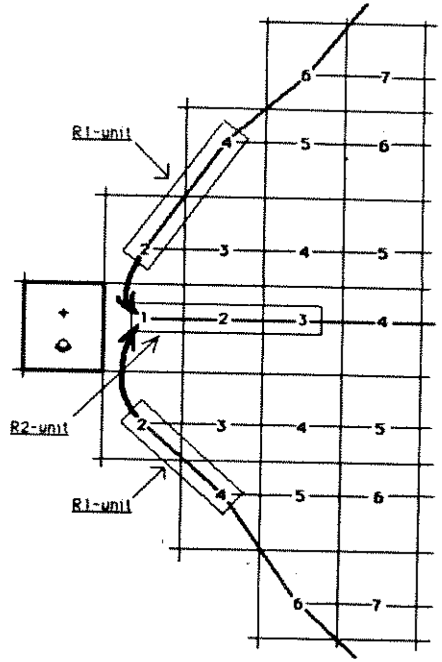


(a)

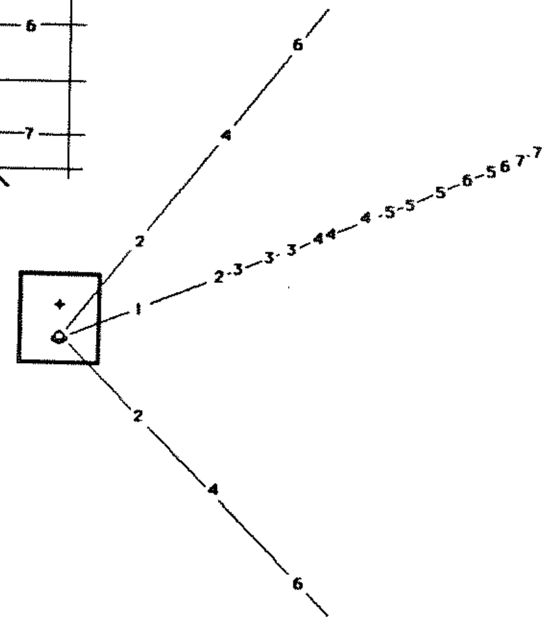


(b)

Figure 5. Reverberation units; (a) R1-unit, (b) R2-unit.



(a)



(b)

Figure 6. Two-dimensional crossfeeding; (a) interaction of R2- and R1-units, (b) the resulting reverberation streams.