International Journal of Multimedia and Ubiquitous Engineering
Vol.8, No.6, (2013), pp.201-208
http:// dx.doi.org/10.14257/ijmue.2013.8.6.20

Audio Analysis & Creating Reverberation Plug-ins In Digital Audio Engineering

Yoemun Yun

Dept. of Applied Music, Chungwoon University
Daehakgil-25, Hongseong, Chungnam, South Korea

hippie740@chungwoon.ac.kr

Abstract

In audio engineering, early reflection patterns are not usually variable, such as the user makes a selection from a handful of stored patterns simulating various rooms, halls, chambers and plates, or small room ambience, reflections is variable. To create digital reverb is much more complicated than combining a few digital relay units. The greater the spacing the larger the room sounds. Some difficult problems happened in recording situation is that every acoustic instrument engineers record sounds small or dry because their recording room does not have the same degree and circumstances of acoustic ambience as good as concert halls or venues. As a matter of fact that the same situation is also true of the most professional studios, apart from those with large, live rooms, and most pop music is recorded dry enough, either in a dry environment. In that case, people use close miking technique to get right sound source. Electronic effects units are used to recreate the desired ambience, and by far the most important effects unit of them all is reverberation.

Keywords: *Reverberation, RT60, Sabins, MAX/MSP, Sound Reflection*

1. Introduction

In real world, reverberation happens naturally when sound is bounced and re-reflected from walls, ceilings and other surfaces within a large room, and an electronic reverb unit imitates this effect by generating thousands of reflections electronically. Reverb can be used to create the impression of a real room, but it may also be used to enhance new effects that have no absolute counterpart in nature.

If engineer record in a dry environment and then add some effects during a mix, engineer can have experience with different types of effect after the recording has been done. This is practical situation in the studio with experienced mix engineers, many tracks and variety different effects units.

For recorded singing voice, however, it is almost always accompanied by musical instruments that in most cases are harmonic, broadband, and are correlated with singing since they are composed to be a coherent whole with the singing voice [1].

Numerous reflections happen in a room because the sound emitted from a source hits the surfaces of the room at an infinite number of angles. Sound hits surfaces and reflections until all of energy is absorbed and disappear.

The first sound is called direct sound and the next comes known as early reflections [2]. The reflections allow the sound to bounce for a short time until it is absorbed. This characteristic of sound is called reverberation or reverb. These reflections also spread the sound throughout the entire room.

ISSN: 1975-0080 IJMUE

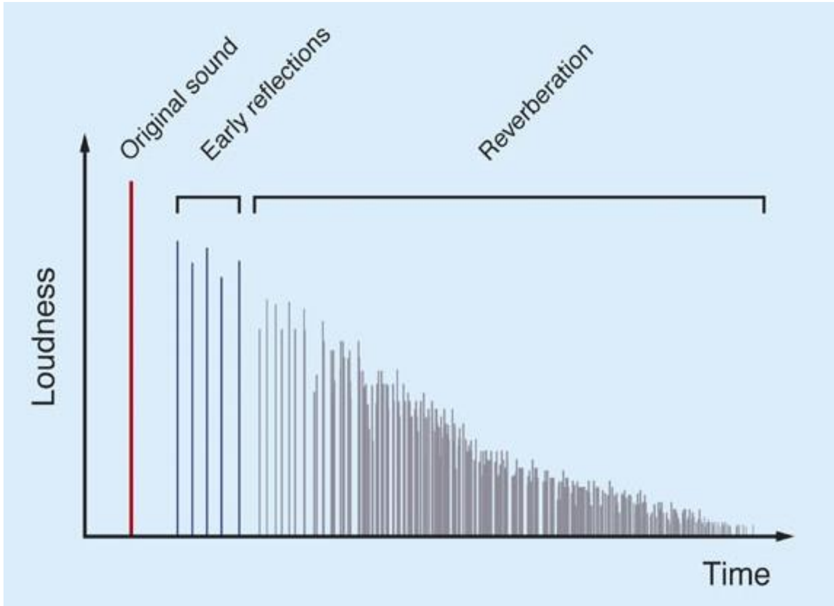

Figure 1. Characteristic of Sound and Reverberation

2. Reverberation

2.1 Reverberation time

Reverberation time (T60) is defined as the time in seconds for a sound to decay by 60 dB after the sound source has stopped. Generally, the reverberation time increases as the room volume increases, and reverberation time decreases as the total amount of absorption increases, as would be expected [3].

Volume and absorption can both be calculated. Reverberation times may be calculated from volume and absorption. The absorption of most materials is frequency dependent. As a direct result, the reverberation times in rooms are frequency dependent. Consequently, the total absorption in a room must be calculated separately at each frequency band of interest.

The amount of absorption provided by a material (at a given frequency) is the surface area of that material multiplied by its absorption coefficient (at the given frequency). The total absorption in a room (at a given frequency) is the sum of the absorption (at the given frequency) of each of the materials in the room.

$$A = \sum S\alpha = \sum a$$

$$\sum a = S_1\alpha_1 + S_2\alpha_2 \ldots S_n\alpha_n$$

A = Total absorption in sabins (square feet)
S = Surface area of each material in square feet
α = absorption coefficient
a = absorption of each material type in sabins

The Sabine Equation is the most commonly used of the reverberation equations because it generates useful results in a wide variety of spaces. Absorption coefficients are used in the calculation of reverberation times in rooms. The absorption of any materials is measured at many frequencies to determine their absorption characteristics [4].

Reverberation times at each frequency band of interest may be calculated as follow.

$$T_{60} = \frac{.05V}{A}$$

T60 = reverberation time in seconds
V = volume of the room cubic feet
A = total absorption in sabins
.05 = constant in seconds per foot

Although speech separation has been extensively studied, few studies are devoted to separating singing voice from music accompaniment. Singing voice bears many similarities to speech.

For example, they both consist of voiced and unvoiced sounds. But the differences between singing and speech are also significant. A well known difference is the presence of an additional formant, called the singing formant, in the frequency range of 2000-3000 Hz in operatic singing [1]. It must be pointed out that sound fields in rooms are exceedingly complex, and some simplifying assumptions must be made in order to derive manageable equations.

For example, the Sabine equation assumes that the sound field is evenly distributed throughout the room, and neglects the specific locations of various absorptive materials as well as the shape of the room.

Moreover, although T60 is a fundamental room descriptor, other considerations such as initial time delay, spatial impression, blend, and freedom from intrusive noise may also be important.

2.2 Digital reverb

Almost virtual digital reverbs use a single reverb time for the entire range of frequencies, as well as a single algorithm. Some models can vary reverb time with frequency and some can equalize the input which may enhance the sound by changing tone color.

The base of idea is the audio's key characteristics that many audio streams can be mixed into the stream equivalent to just one stream in size. When various sounds are audible, we can hear some of the loudest sounds. This is because the sounds having different frequency counterbalance each other. With other media such as video, text, we can't see these mixing effects [5].

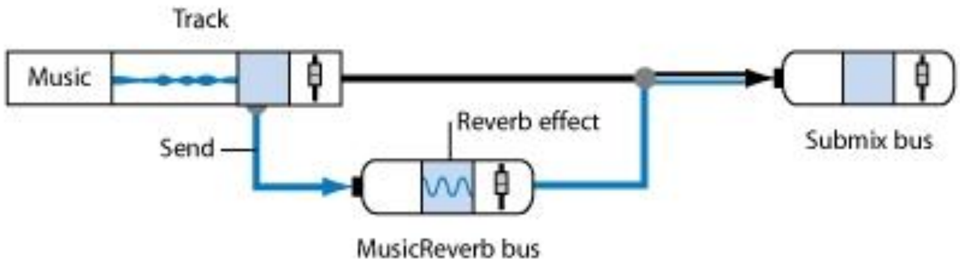

Figure 2. Basic Reverb Signal

To create digital reverb is much more complicated than combining a few digital relay units. Many researches have shown that somewhere between 1000 and 3000 separate echoes are needed every second to create the illusion of dense, natural sound reverberation. [6] and [7] also, the distance between these reflections has to be chosen very carefully or the resulting reverberation will ring in a most unnatural ways.

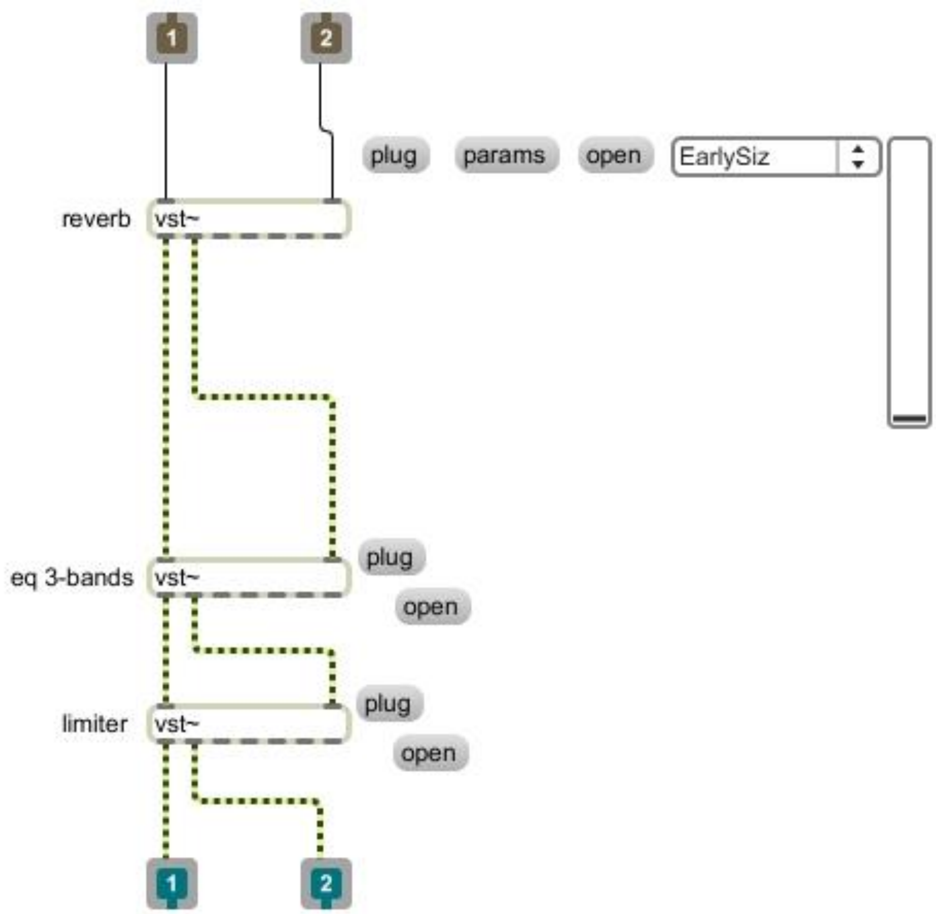

Figure 3. Basic Reverberation Programming in MAX/MSP

Instead of breaking a signal down into individual samples, the parity coding method breaks a signal down into separate regions of samples and encodes each bit from the secret message in a sample region's parity bit [8].

Getting appropriate reflections is only part of the story. Digital filtering is needed to impart the specific frequency characteristics to the reverberant sound, and the software needs to be sophisticated enough to create a number of different simulated environments without becoming too complicated for the users to set it up.

However, digital system allows users to create reverbs that are longer and brighter then anything normally came upon in the real recording.

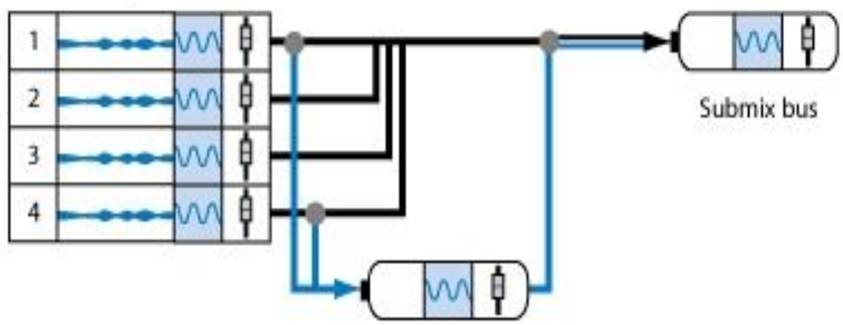

Figure 4. Combining Track Reverb Send

Those models currently available and usually fall into two types [6]. One is the fully programmable variety and the less costly preset based on the machine. It may only be one user-variable parameter per preset.

The basic spread spectrum technique, on the other hand, is designed to encode a stream of information by spreading the encoded data across as much of the frequency spectrum as possible. This allows the signal reception, even if there is interference on some frequencies [8].

The rapid advancement of computer technologies contributed to make the speed of processor. There appeared more Internet-users, protocols and programs requiring considerable throughput.

There has always been and will continue to be some kind of effort to minimize required computer processing capacity and network usage in order to save costs [5]. The human auditory system perceives over a range of power greater than one billion to one and a range of frequencies greater than one thousand to one.

Sensitivity to additive random noise is also acute. The perturbations in a sound file can be detected as low as one part in ten million (80 dB below ambient level) [8]. In a real acoustic environment, speech is usually contaminated by interference that can be harmonic or non-harmonic, narrowband or broadband.

Interference in most cases is independent of speech in the sense that the spectral contents of target speech and interference are uncorrelated.

For recorded singing voice, however, it is almost always accompanied by musical instruments that in most cases are harmonic, broadband, and are correlated with singing since they are composed to be a coherent whole with the singing voice [1].

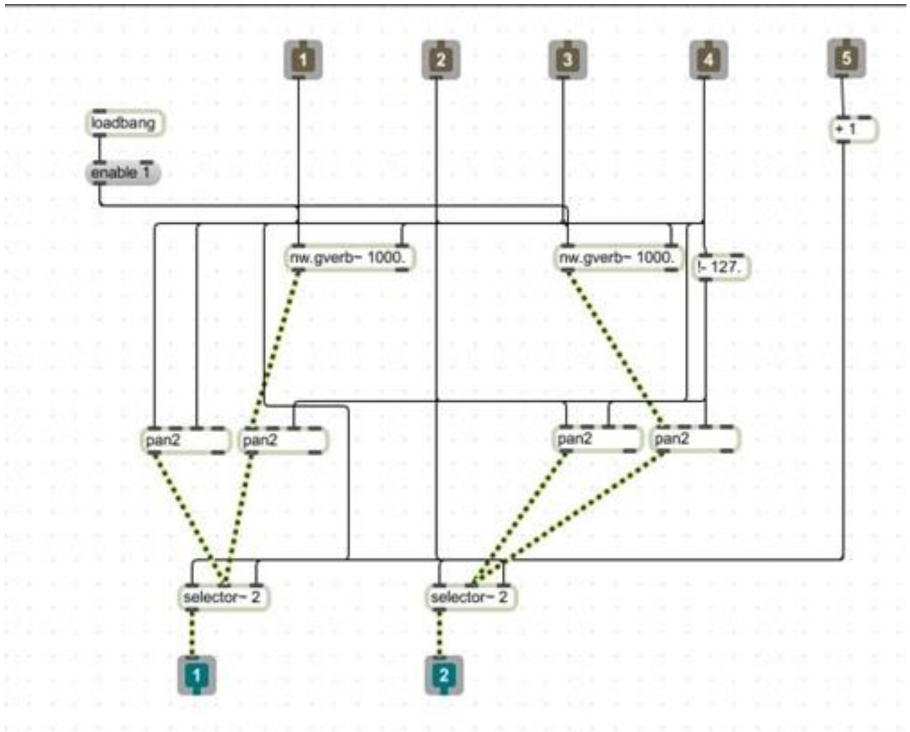

Figure 5. Complicated Reverberation Programming

If audio content having fewer channels than can be provided by a target system is available, the target audio system cannot take full advantage of it. Therefore, in order to utilize such audio content, it is necessary to use an upmixing method that converts mono or stereo audio formats into a multi-channel audio format suitable for such a system [9].

If users test the programmable type it will then be easy to understand the adjustments that must be made in a preset machine. Both types have stereo outputs and may have either mono or stereo inputs, depending on the model's preferences.

However, the reverb in either case is generally derived from a mono mix of the inputs [10]. Only the dry signal remains in stereo throughout. This concept is suitable in practical situations because reverb comes from all directions regardless of the stereo positioning from the original sound and reverb in natural sound a mono in/ stereo out system is all that is needed.

2.3 Available parameters

The main reverb parameters available for user control are pre-delay time, early reflection pattern, level, overall decay time and high-frequency damping. Pre-delay simply consist of the time between the original sound and the first reflection, and may be variable from virtually instaneous to half a second or more.

This is a simple way to create the illusion of room size and also helps to separate the dry sound from the reverb. Longer pre-delays can be useful in combination with medium decay vocal reverb to prevent the reverb from muddy the vocals.

Figure 6. Digital Reverberation

Early reflection patterns are not usually variable, such as the user makes a selection from a handful of stored patterns simulating various rooms, halls, chambers and plates, or small room ambience, reflections is variable. The greater the spacing the larger the room sounds.

A time delay element is used to provide such ambience effects, and a ±90° phase shifter is needed to present spaciousness effects. Assuming that the distance from front loudspeakers to the wall is about 2 meters and the distance from rear loudspeakers to the wall is about 1 meter, a time delay of 12 ms is applied to the surround channels [1].

Overall decay time simply determines how long the reverb takes to disappear, and longer reverb times suggest large environments with very reflective surfaces whereas shorter ones may be used to simulate the natural acoustics of a small room. Most reverb units can produce a long decay time, but the true test of a reverberation is how persuasively it imitates small room ambience [11].

High frequency damping allows the high-frequency decay time to be shorter than the overall decay time in order to simulate the absorbency characteristics of real rooms, simulating the absorption of both surfaces and air. Some units also have independent control over low-frequency damping to simulate environments that reflect mainly high-frequency sounds.

Not the only audio data size, the CPU load in decompressing the compressed audio data should be also considered. The data volume and CPU load from data decompression would be increased rapidly [5]. The input is first partitioned into spectrally homogeneous portions detecting significant spectral changes. Then, from energy comparison between differential signals between stereo AR (all recorded) signal and each channel signals in frequency domain, the presence or absence of singing voice were determined.

The proposed systems were implemented in DSP board. And input AR stereo signal, extracted MR (music recorded) signal from DSP output, and extracted singing voice could be listened in real-time by multichannel audio interface cooperated with DSP board [1].

By selecting the appropriate pattern for the environment to be simulated and then adjusting the other parameters. The available effects can vary from a barely reverberant room to a huge cavern, in which the reverb decay thunders on for several tens of seconds.

In order to utilize such audio content, it is necessary to use an upmixing method that converts mono or stereo audio formats into a multi-channel audio format suitable for such a system [9].

In real, most of the useful applications of reverb have a decay time below two seconds because excessively long reverb tends to muddy a mix although there are circumstances in which a long reverb can be effective.

3. Conclusion

Some digital reverb models incorporate a control which governs the reverb's room size by adjusting several of the parameters at the same time to give the impression of large or smaller spaces.

Singing voice removal algorithm consists of power comparison between each channel of the stereo signal and inter-channels differential signal and then spectrally removed at each singing voice frequency bin when larger than threshold value respectively [1]. This is useful because it means users do not have re-programmed several parameters. Reverberation can be applied to many of related virtual effects, such as gate reverb and reverse reverb. Also, delay, echo, flanger and chorus can be made by tweaking reverberation.

References

[1] H. T. Kim, T. H. Kim and J. S. Park, "A Real Time Singing Voice Removal System Using DSP and Multichannel Audio Interface", International Journal of Multimedia and Ubiquitous Engineering, vol. 7, no. 2, (**2012**).
[2] D. E. Hall, "Musical Acoustics", Third Edition, Books/Cole, (**2002**).
[3] K. A. Hoover, "An Appreciation of Acoustics", Cavanaugh Tocci Publishing, (**1991**).
[4] G. D. White, "The Audio Dictionary", Second Edition, University of Washington.
[5] Y. M. Kim, "The Audio Control using TDM (Tree-structural Distributed-terminal Mixing) in the Webcast", International Journal of Multimedia and Ubiquitous Engineering, vol. 1, no. 4, (**2006**).
[6] http://manuals.info.apple.com/en/soundtrack_pro_user_manual.pdf.
[7] F. Rumsey and T. McCormick, "Sound and Recording", Fourth Edition, Focal Press, (**2002**).
[8] P. Dutta, D. Bhattacharyya and T. Kim, "Data Hiding in Audio Signal: A Review", International Journal of Database Theory and Application, vol. 2, no. 2, (**2009**).
[9] C. J. Chun, Y. G. Kim, J. Y. Yang and H. K. Kim, "Real-Time Conversion of Stereo Audio to 5.1 Channel Audio for Providing Realistic Sounds", International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 2, no. 4, (**2009**).
[10] www.cycling74.com, MAX/MSP Manual, **(2009)**.
[11] J. Chappell, "Digital Home Recording", Backbeat Books, (**2003**).

Author

Yoemun Yun

He received his B.S. degree in Music Production & Engineering from Berklee College of Music, Boston, MA, USA in 2004, and his M.S. degree in Music Technology from New York University, NY, USA in 2006. He is now an Assistant Professor in the Department of Applied Music, Chungwoon University, Chungnam, Korea. His research interests include music engineering, audio sound design, analog-to-digital audio conversion, electronic music, and computer music composition.

