DESC9115 Digital Audio Systems Final Written Review

Name: Zhe Shi

SID: 450105812

Implement Artificial Reverberation by using filterbank methods

Abstract

This final review will present the digital implementation of artificial reverberation by using filterbank methods. It will mainly focus on discuss and analysis the basic digital audio principle of each component as well as the overall structure of the classical algorithmic reverberator. Sound examples will also be provided to show how reverberator produce a room support or tone color change for different instrument sounds and vocal sound.

Problem description

Reverberation is one of the most important characteristics in room acoustic quality. It is also an important component when describing tone color or tone quality of the music. The reverberation can be defined as "the persistence of sound after the original sound is produced."[2, P425-426]. It determined by original sound bounced and reflected by the boundary and shape of the room. In music, a relatively longer reverberation will provide more "warmth" and "fullness" sensation. Human's voice would also benefit from a certain level of room support as well. Nevertheless, if reverberation is too long, the intelligibility of the vocal sound would decreasing. In order to measure the reverberation time, the most common way is to use an impulse response and the sound of the pistol is the most frequently-used way to create impulse response in practical. The reverberation will be treat as two separate parts: early reflection and late reverberation, both served a different purpose when defining the acoustic qualities in a room.

In digital signal processing, there are two ways to produce artificial reverberation: convolution and filterbank. Convolution method is to use the impulse response that recorded in the real rooms or actual halls and then convolutes the input signal and the impulse response to create a relatively more realistic and faithful reverberation. Filterbank method, on the other hand, is to use algorithm to simulating a room model and create a imitate version of real reverberation. Since the filterbank method is a simulation, it is harder to create a natural reverberation as convolution method. However, filterbank method provides an option to doing some detailed parametric changes, which is hard for convolution method to achieve.

Specification

The most fundamental element of the filterbank reverberator structure is the decay line. Schroeder has made lots of efforts to create the artificial reverberation by using multiple delay line at the early 1960s.[1] He uses a series of allpass filters to mimic the real reverberation or in his words called the "colorless artificial reverberation." Unlike low-pass and high-pass filter, Allpass filter does not change the amplitude of the signal at all frequency, but it will alter the phase of the signal and create decaying echoes.[1] The early all-pass filter series shown as [figure 1]

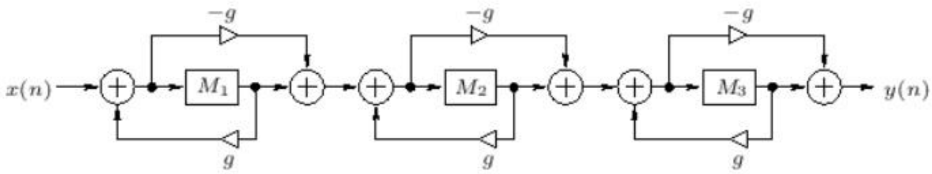

figure 1 the Schroeder's allpass filter series[2]

In Schroeder's design, the decay factor g is set at 0.7. and the delay time M are series of prime numbers (such as 241,757,113).

Another critical component for the filterbank reverberator is the feedback infinite impulse response (IIR) comb filter.[1] As the name indicated, comb filter has an appearance of comb in the frequency response. It will add a delayed edition of the original signal to itself. [figure 2]

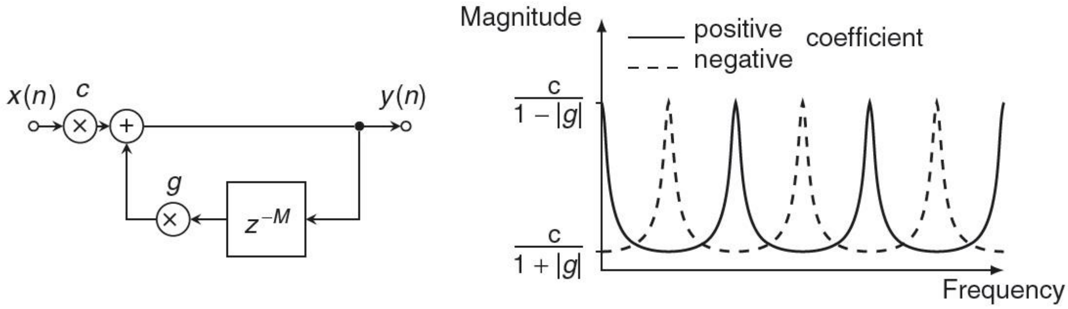

figure 2 IIR comb filter and magnitude response[1]

As the figure 2 indicated, the delayed version the of signal from output will feedback to the input. It will be implemented by the difference equation:

y(n) = cx(n) - gy(n-M)

Since comb filter is an Infinite impulse response (IIR) filter, it is essential to keep the feedback coefficient g less than 1, or the echoes will sound louder and louder and create a never stop sound. Due to the air absorption, the high-frequency part of the natural reverberation contain less sound energy, but the comb filter does not have any treatment to simulate this particular phenomenon. [1, P99]. A quick solution for this problem is to add a first order low pass filter into the comb filter. (figure 3) to attenuates the output signal the at high frequency.

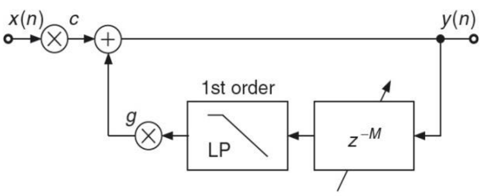

figure 3 Low-pass IIR comb filter[1]

Despite the allpass filters and comb filter can successfully create a sense of decaying echoes, it is only good to manipulating the late reflection of the artificial reverberation. The early reflection, on the other hand, is more relevant and correlated to more vital factors of the room acoustic qualities and the characteristics. In a narrow sense, early reflection can be defined as sound received within roughly 100 milliseconds after the direct sound signal arrived. It will influence the sense of spatialness (listener's spatial impression) and instrument's timbre. In order to further adjusting the early reflection in the artificial reverberation, taped delay line have been introduced into the filterbank reverberators. Tapped delay line is delay line with multiple "taps" or "reading points" that can extract signal from the delay line, each extracted signal will be weighted and summed up to create an output signal finally. According to Udo Zolzer:"it was noticed that the early reflections have great importance in the perception of the acoustic space and that a direct-form FIR filter can reproduce these early reflections explicitly and accurately. "[1, P168] The general causal Finite Impulse Response (FIR) filter implement tapped delay line (Figure 4).

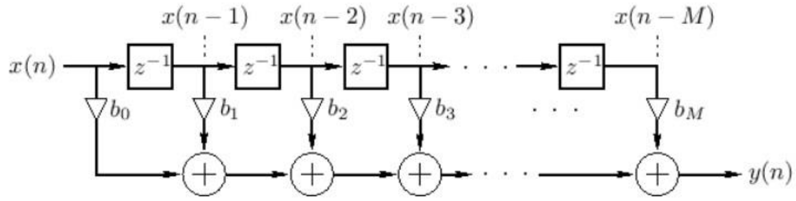

figure 4 The general, causal, finite-impulse-response (FIR) filter [1]

Implementation

As mentioned before, early reflection can define people's perception of the auditory space shape, and it should take care separately from the late reflection. The structure of algorithmic reverberation would look like this (figure 5)

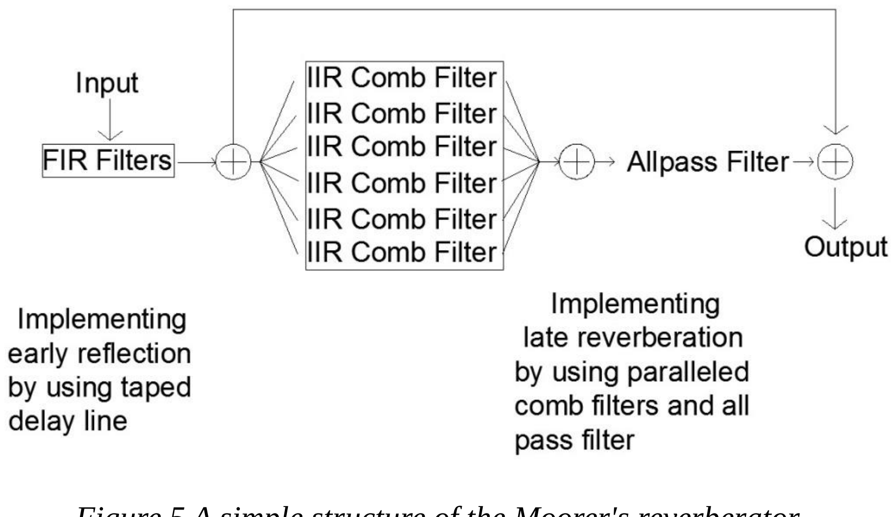

Figure 5 A simple structure of the Moorer's reverberator

Based on the Schroeder's classical reverberator, Moorer has made a noticeable improvement in his version. As the figure 5 indicated, the input signal will go through the tapped delay line to deal with the early reflection at first, and then the summed result from the implementation of early reflection will enter the late reverberation section. This part consist of six parallel comb filters and an allpass filter, every comb filter has a low-pass filter in it to simulating the nature reverberation and cancel the unwanted high-frequency sound in the reverberation. Each comb filter has a decay coefficient g, which controls the overall decay time of the reverberation. In order to acquire the desired decay time, the gain for each tap should be set to

$$g_i = 10^{-3\frac{T_d F_s}{m_i}}$$

Where Td is the desired decay time, Fs is the sample rate, and mi indicates the length of delay in samples. The decay coefficient for the lowpass filter also link to the particular frequency's decay time, further experiment and tuning is required to get the best value for it. In Moorer's original works, he provides a recommended ratio between the feedback decay coefficients in the comb filter and the coefficients in allpass filter. Those two values should be maintaining at the range 50 to 80 milliseconds and have a 1:1.5 ratio. He also points out that the reverberator would have better results if the taps can be located related to the computed geometric modeling techniques.[1, P187]. After the late reverberation treatment, the early reflection, and late reverberation will add together and produce the final output signal.

Evaluation

Vocal enhancement

An echo-ish long reverberation will reduce the intelligibility of the vocal sounds significantly and may also make a rough and robotic coloration on the original sound. In order to provide a relatively more natural reverberation sound for human voice, it is essential to minimize the modification induced by the later reverberation and more focus on the early reflection part.

The input signal "Densil1.wav" is the vocal sound which recorded in an anechoic chamber. "Densil2.wav" is the output that only use the taped delay line section of the reverberator, the vocal sound will sounds warmer (have little bit more bass) and the sound image is slightly more close to the listener's ear level. However, if the input sound keep going through the comb filter and all-pass filter parts, ("Densil3") output will sound blur and unclear.

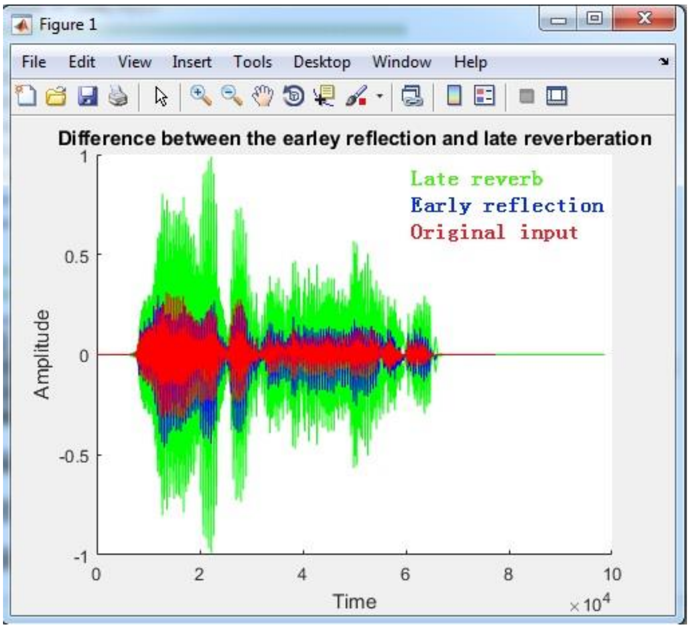

Figure 6 The difference between the early reflection and late reverberation - vocal sound

That is to say, if the purpose is not for anti-mainstream artistic approaches, restrict the changes on reverberation in a small range is necessary for the vocal sound enhancement process. Moreover, intelligibility and reverberation are two totally opposite direction, finding the balance between this two characteristic could be a vital key for vocal enhancement.

Instrument sounds enhancement

Unlike the vocal sound, instrument note's may require a relatively longer reverberation time. The input signal "1.wav" is a single xylophone note. "2.wav" is the input signal only go through early reflection part and "3.wav" is the signal have both early reflection and late reverberation. By carefully tuning the decay length and factors, both the early reflection and the full reverberation is sounded natural, convincing and with more fullness and warmth than original sound.

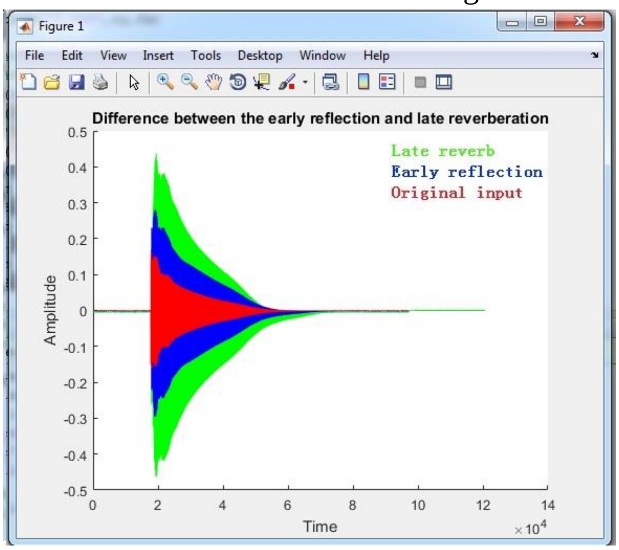

Figure 7 The difference between the early reflection and late reverberation - Xylophone note

The "erhu.wav" is another input signal for testing how early reflection and how the reverberation will affect the overall tone color and tone quality of music. Erhu is a traditional Chinese instrument, and it is like a violin but only have two strings. The input signal is a segment cut from an Erhu solo.
After different test sets of parameters, this string music is founded not as sensitive as previous percussion sound when modifying the reverberation, "erhu2.wav" is the input signal only goes through taped delay line part. the timbre of sound is more open and have slightly more shrillness color at high frequency. "erhu3.wav" is the input signal go through the complete reverberator, by increase the wet mix rate, the sound is more blur and have a bit scratchy noise. As the Figure 8 indicated, the tone color is very delicate and require a careful micro fine tuning for all parameters.

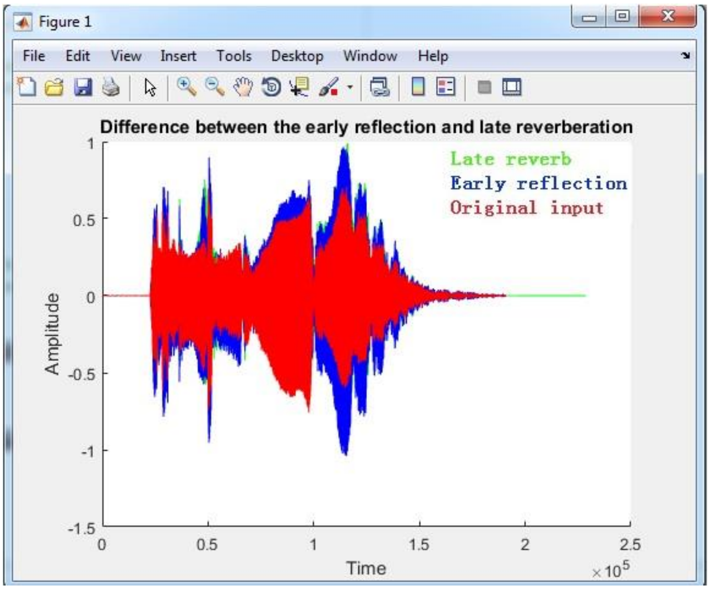

Figure 8 The difference between the early reflection and late reverberation - The Erhu music

To sum up, a good filterbank reverberator is a perfect combination of science and art. It could be time-consuming to find the best parameters for different instruments and music. The vocal sound is more delicate than music and may need some special treatment in order to provide a more natural reverberation.

Reference

[1] Zölzer, U., & Arfib, D. (Eds.). (2011). *DAFX: digital audio effects* (Vol. 1). Wiley.
[2] Valente, M., Hosford-Dunn, H., & Roeser, R. J. (Eds.). (2008). *Audiology: treatment* (Vol. 2). Thieme.