ALL NATURAL ROOM ENHANCEMENT

Kyle Spratt, Jonathan S. Abel

Stanford University
Center for Computer Research in Music and Acoustics (CCRMA)
Stanford, CA, USA
kspratt@ccrma.stanford.edu

ABSTRACT

A recording technique due to Walter Murch for extending the reverberation time of a room is analyzed, and a real-time implementation is presented. The technique involves speeding up a prerecorded dry sound and playing it into a room. The room response is recorded and subsequently slowed down such that the original signal appears at its normal speed, while the reverberation of the room is 'stretched,' causing the room to sound larger than it is. A signal analysis is presented showing that this process is equivalent to slowing down the room impulse response. Experimental measurements on a simple physical system confirm this effect, and show that the process can be interpreted as either scaling the room dimensions, or slowing the sound speed. Finally, we describe a block processing approach which implements this technique in real time with a fixed processing latency.

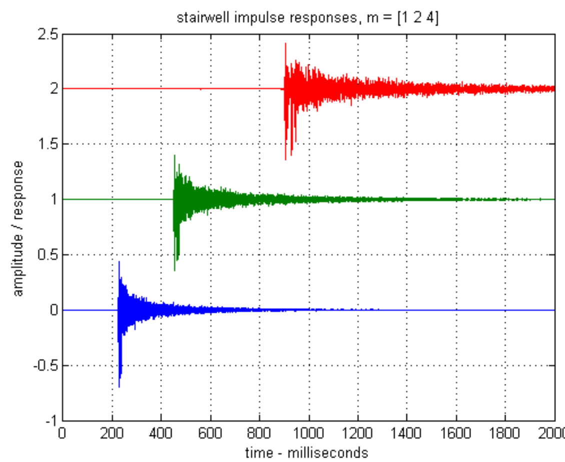

Figure 1. Example impulse responses.

1. INTRODUCTION

This paper is concerned with a recording technique used to modify the reverberation time of an acoustic space. The technique, discovered by Walter Murch while working as a sound editor for motion pictures in the late 1960s [1], is best described in terms of the analog tape machines that were used for recording at the time: One starts with a tape recording of a dry sound which has been recorded at normal speed. This tape is played into the room at a faster tape speed than it was recorded. Using a second tape machine, the sped up sound is recorded at the same, faster tape speed at which the source is being played. When this second recording is then played at the original recording speed, the source signal is returned to its normal speed, while the reverberation added to the signal by the room is 'stretched,' causing the room to sound larger than it is. As an example, Figure 1 shows the technique applied to a small room, stretching its response to a short pulse by various factors.

In the sequel, we explore this method, analyzing its signal flow as well as presenting measurements made using a relatively simple physical space. Analogies are drawn both in terms of signal processing concepts, and also in terms of the physical characteristics of the room. Finally, we describe a digital implementation of the technique which can be done in real time with a fixed amount of processing latency.

2. ANALYSIS OF METHOD

2.1. Signal Analysis

We denote our input signal as $s(t)$. We are assuming that the room is acting as a linear and time-invariant system, such that playing the signal into the room and recording it at some point can be described as convolving the input signal with the impulse response of the room $h(t)$,

$$(s * h)(t) \doteq \int_{-\infty}^{+\infty} s(u) \cdot h(t-u) du. \tag{1}$$

Speeding up a signal can be described mathematically as a contraction of the time axis, and as such the input signal sped up by a factor m is given by

$$\acute{s}(t) \doteq s(m \cdot t), \tag{2}$$

where $\acute{s}(t)$ denotes the sped-up signal. The response of the room, $\acute{r}(t)$, to the sped up signal is then

$$\acute{r}(t) \doteq (\acute{s} * h)(t) = \int_{-\infty}^{+\infty} \acute{s}(u) \cdot h(t-u) du. \tag{3}$$

Slowing down $\acute{r}$ to the original speed is a dilation by m, and hence the result $r(t)$ of the entire process is

$$r(t) \doteq \acute{r}(t/m) = \int \acute{s}(u) \cdot h(t/m - u)du. \tag{4}$$

Changing the variable of integration from u to u/m yields

$$r(t) = \int \acute{s}(u/m) \cdot h(t/m - u/m)du/m$$

$$= \int s(u) \cdot \frac{1}{m} h\left(\frac{t-u}{m}\right) du = (s * \grave{h})(t). \tag{5}$$

Where we noticed that $\acute{s}(u/m) = s(m \cdot u/m) = s(u)$ and we have denoted

$$\grave{h}(t) \doteq \frac{1}{m} h\left(\frac{t}{m}\right). \tag{6}$$

We conclude that the effect of the entire process is that we have convolved our original input signal with a scaled and time-dilated version of the natural impulse response of the room, where both the scaling and the dilation are by the same factor m used to contract the original signal.

It's worth mentioning that in the above analysis we did not restrict m to be greater than 1. This implies that the technique can also be used to shorten the reverberation time of a room, by slowing down the original input signal and then speeding up the subsequent recording.

2.2. Experimental Measurements

In order to test the conclusion of our signal analysis, we measured the effect of the technique on a relatively simple acoustic space, namely an acoustic tube with both ends closed. The tube was made of ABS plastic, 4" in diameter and roughly 10' long, with the speaker enclosing one end and a rigid plastic cap on the other, the microphone—2mm in diameter—being placed flush against the cap on the inside of the tube.

We used an exponentially swept sine signal as input, so as to measure an impulse response for the system [2]. The result is expected to be a sequence of equally spaced decaying pulses, with the spacing λ equal to

$$\lambda = \frac{2 \cdot L}{c}, \tag{7}$$

where L is the length of the tube and c is the speed of sound. This was the case for the natural impulse response of the tube (see Figure 2). For the cases when the scaling factor m was 2 and 4, we see that the spacing between the pulses was $2 \cdot \lambda$ and $4 \cdot \lambda$, respectively.

This leads us to conclude that using a time-scaling factor of m will result is the pulses being spaced by $m \cdot \lambda$, which is in agreement with the signal analysis conclusion that the impulse response will be dilated in time by the same factor used to contract the input signal. In terms of the above physical equation, this implies that

$$m \cdot \lambda = \frac{2 \cdot (m \cdot L)}{c} = \frac{2 \cdot L}{c/m}, \tag{8}$$

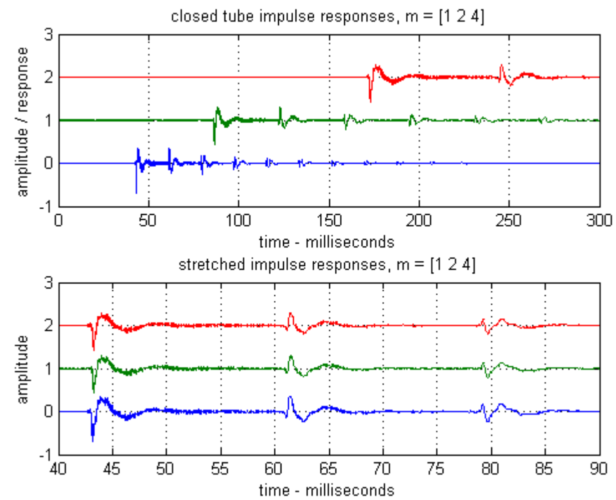

Figure 2. Tube impulse responses for $m = 1, 2, 3$.

which gives us a pair of physical analogies for the effect of using the technique. According to the above equation, we can interpret the effect as being equivalent to either scaling the length of the tube by m, or scaling the speed of sound by $1/m$.

In terms of a normal room, we can extend the first interpretation by assuming that we have scaled all of the dimensions of the room[1] by m, hence scaling the volume of the room by m^3. Since the wavelengths of the dry sound have been scaled by $1/m$, by scaling the physical dimensions of the space we have returned the sound waves back to their original frequency. W. C. Sabine applied a similar technique during his pioneering studies of reverberation in the beginning of the 20^{th} century, using scale models of auditoriums excited by ultrasonic sound waves in order to study the subsequent patterns of reflections [3].

For the second interpretation, assuming that the speed of sound has been decreased is equivalent to assuming that the density of air has increased. In fact, in the late 1950s a 'miniature reverberation chamber' was invented using this principal [4], wherein a small chamber containing a speaker and microphone was filled with a gas much denser than air (such as freon or vaporized mercury). The density of the gas implied a slower speed of sound, and this caused the chamber to sound larger than it was.

2.3. Effect on Late Reverberation

It has become customary to distinguish between the early portion of a room's impulse response, made up of a sparse set of 'early reflections,' and the later portion of the response, which is generally considered to be an exponentially

[1]Note that all objects in the room, including the speaker and microphone, are also scaled. This scaling will affect, for instance, the radiation pattern of the speaker, the polar pattern of the microphone, and the scattering of sound by objects in the room.

decaying diffuse field [5]. From the results obtained thus far, it seems reasonable to conclude that when using this recording technique, the early reflections will become dilated in time, perceptually causing the space to sound larger.

When considering the late reverberation tail, since it is a diffuse field rather than a collection of discrete arrivals, it might be imprudent of us to make the same assumptions. The τ_{60} of the reverberation, defined as the time it takes the response to decay by 60 dB, is a classic way of characterizing the reverberation of a room. Since we are stretching the time axis of the room's impulse response, it follows that we are also scaling the τ_{60} by the same factor.

However, this line of reasoning ignores the frequency-dependent characteristics of the late reverberation. A more accurate description of the reverberation tail would be given by considering the rate at which different frequencies decay in the room, which can be seen by viewing the spectrogram of the impulse response. This amounts to making τ_{60} a function of frequency, hence $\tau_{60}(f)$ for the frequency f in Hz.

We measured the impulse response of a stairwell when using various scaling factors in order to see how the late reverberation is affected. Figure 3 shows impulse response spectrograms for the various scaling factors. From the upper plot, it appears that the spectrograms are being both 'squashed' by a factor m along the frequency axis, while being 'stretched' by a factor m along the time axis. The lower plot, showing a rescaling of the spectrograms, confirms this.

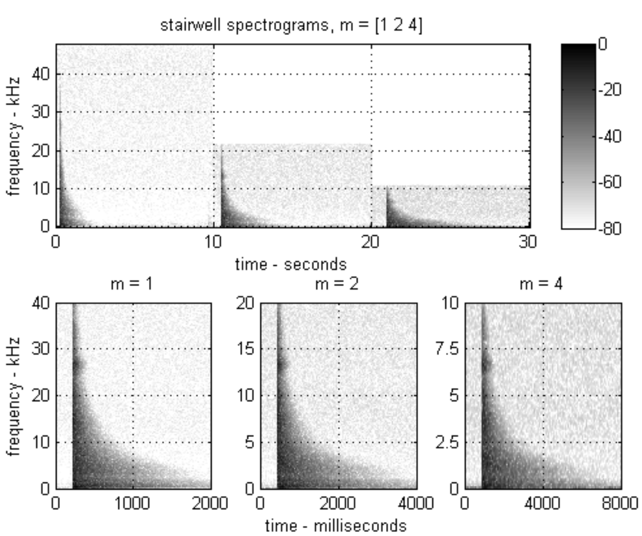

Figure 3. Stairwell response spectrograms.

Using the same responses, Figure 4 shows a plot of τ_{60} as a function of frequency, where both axis have been plotted logarithmically. For the natural response of the room, the lowest of the plots, we see that the τ_{60} rises with frequency until reaching a peak in the 700 Hz range, and then decreases. When using the technique, the general shape of the curve is unchanged, save for being translated along a diagonal, with frequency decreasing and τ_{60} increasing.

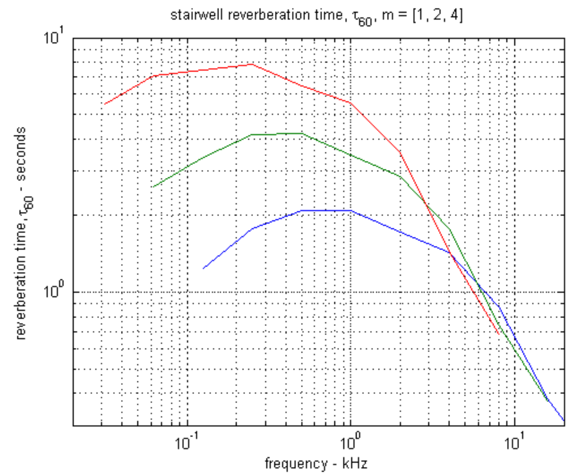

Figure 4. Reverberation time as a function of frequency for various scaling factors.

3. REAL-TIME IMPLEMENTATION

The previous descriptions of the technique, based on the existence of an entirely prerecorded dry signal, seems to indicate that it is an inherently non-real-time process. However, by using the signal processing methods described below we are able to apply the technique to a live input signal in real time, with a fixed amount of processing latency.

Figure 5 shows a signal-flow diagram of the implementation. For our particular implementation, we require that the scaling factor $m \geq 1$ be an integer. We use down-sampling (dropping samples) and up-sampling (inserting zeros between samples) as a means to contract and dilate the discrete-time signal, respectively, and as such require filtering to avoid aliasing and imaging. An important consequence of this filtering is that we lose a fair amount of bandwidth through the entire process. If we assume that our system is running at a sampling rate f_s, then the bandwidth of our output will be reduced to $f_s/2m$; see for example the spectrograms in Figure 3.

The signal flow can be broken down into two separate sections, with the actual physical room acting as the bridge between the sections. The first section takes the original dry signal as input, applies an anti-aliasing filter and down-samples the signal by the scaling factor m, then buffers and zero pads the signal with enough space for the room to decay naturally. The zero-padded buffer is then played into the room. Though the down-sampled signal is being played faster than the original, and hence lasts a shorter amount of time, the zero padding allows enough time for another buffer of down-sampled data to be stored by the time it is needed to be played into the room.

It should be noted that this implementation only allows

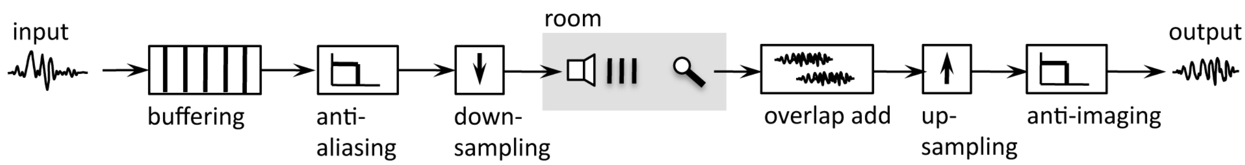

Figure 5. Real-time implementation signal-flow architecture.

for the lengthening of reverberation time, caused by speeding up the original sound. Shortening the reverberation time ($m < 1$) would require slowing down the dry signal and then playing it into the room. Our real-time implementation relies on the fact that when down-sampling the signal, we can then zero pad in order to let the input signal 'catch up' with the output. This would not be possible if we were slowing down the dry signal instead of speeding it up.

In the second section of our signal flow, the input is the sound of the room, which is then put into a buffer using an overlap-add method [6], up-sampled by a factor m and put through an anti-imaging filter (which is identical to the anti-aliasing filter used in the first section). The output is then equivalent to performing the recording technique using the non-real-time method described above.

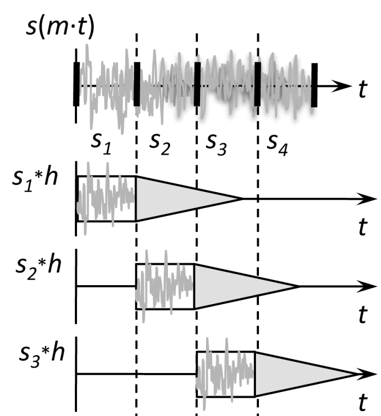

Figure 6. Overlap-add response reconstruction.

Figure 6 is a graphical representation of the overlap-add buffering that we use to reconstruct the room response. Since we are up-sampling the buffered data before outputting it, the data is coming into the buffer faster than it is going out. In the time it takes to output one block of sped-up data, the entire reverberation tail will have been buffered. This allows us to 'double back,' and to overlap the next incoming buffer of data starting at the end of the last block (in other words, overlap with the last reverberation tail). In this way, we output the blocks of sound in a continuous, uninterrupted fashion, while also retaining the entire reverberation tail of the signal blocks.

While there is an inherent processing latency involved in this implementation, the amount of latency can be reduced using parallel processing, although to be done transparently this would require multiple copies of the space being used which are physically identical to each other while being acoustically independent from one another. This could be done using objects which are mass produced, such as metal trash cans. Another possibility would be to use identical hotel rooms.

4. REFERENCES

[1] Murch, Walter, interview. **THX-1138**. DVD, Warner Bros. Entertainment Inc., 2004.

[2] Farina, Angelo. "Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique." 108^{th} Audio Engineering Society Convention, February 2000.

[3] Sabine, Wallace C. "Theatre Acoustics," in **Collected Papers on Acoustics**, Dover Publications Inc., New York, 1964.

[4] Rettinger, Michael and Shipman, Carl N. **Miniature Reverberation Chamber**. U.S. Patent No. 2,986,228. May 1961.

[5] Gardner, William G. "Reverberation Algorithms," in **Applications of Digital Signal Processing to Audio and Acoustics**, eds. M. Kahrs and K. Brandenburg, Kluwer Academic Publishers, Norwell, MA, 1998.

[6] Oppenheim, Alan V. and Schafer, Ronald W. **Digital Signal Processing**. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1975.