ROOM IMPULSE RESPONSE SHAPING FOR ENHANCEMENT OF PERCEIVED SPACIOUSNESS AND AUDITORY DISTANCE

Dae-young Jang

Broadcasting Media Research Group
ETRI, Daejeon, Korea
dyjang@etri.re.kr

Han-gil Moon

Applied Acoustic Lab.
Seoul National Univ., Seoul, Korea
fullmoon@acoustics.snu.ac.kr

Kyeongok Kang

Broadcasting Media Research Group
ETRI, Daejeon, Korea
kokang@etri.re.kr

Hoe-Kyung Jung

Dept. of Computer Engineering
Paichai Univv., Daejeon, Korea
hkjung@pcu.ac.kr

ABSTRACT

Room impulse response is one of most important information with localization of sound source in 3D audio. Actually we can adjust the distance and spaciousness of a sound source with impulse response of the room. Through consequent experiments, we found that C80 and EDT are varying systematically with sound source distances, and these variations are due to early reflection decay curves. This paper contains brief explanation of the two parameters as auditory distance cues, shaping of early reflection decay curves for control of auditory distance, and psychological test results of auditory distance control with early refection decay curve shaping. With these validations, we can confirm early reflection decay curve shape is effective factor for control of perceptual auditory distance and spaciousness in the room.

1. INTRODUCTION

People who research on 3-dimensional sound have mainly performed the 3-dimensional sound localization (elevation and azimuth) of the sound sources for the past tens years. But there have been also little researches for the distance cue of the sound source. Many previous researches have revealed that the conventional distance cues are composed of loudness, spectral information, reverberation and binaural information [1], [2].

Distance cues are classified into relative cues and absolute cues. Relative cues include loudness and spectral information of the source. They have an advantage of easy implementation, but also have a limit of their effect, because they cannot determine the absolute distance without priori information about loudness or spectrum of the sound source. Basically loudness cue is according to the inverse square law. And this cue has a very powerful ability to determine changes in the distance of a constant amplitude sound source. Molecular absorption of the air or air absorption filtering on the sound propagation path is the major origin of spectral cue. In fact, over very long distance this cue has a considerable low-pass filtering effect.

On the other hand, absolute cues are consisted of familiarity, binaural information and reverberation. If the listener is sufficiently familiar with the sound sources, the relative cues can be used to judge absolute distance. Listener's familiarity with both the source signals and the acoustic environment is clearly a key factor in any model for auditory distance perception. For near field listening, we can consider IID & ITD as other cues providing not only directional information but also distance information.

Another important distance cue is the relative loudness of reverberation. When sound is produced in a reverberant space, the associated reverberation may often be perceived as a background ambience, separate from the foreground sound. The foreground sound consists largely of the sound that propagates directly from the sound source to the listener, this so-called direct sound decreases in amplitude as the distance to the listener increases. The amplitude of the reverberation, on the other hand, does not decrease considerably with increasing distance. The ratio of the direct to reverberant amplitude is greater with nearby objects than it is with distant objects. Thus, distant objects sound more reverberant than close objects. An example of this relationship used by Wave Arts Acoustic Environment Modeling is diagrammed in Figure 1. The direct sound amplitude drops 6 dB for each doubling of distance. However, the reverberation amplitude drops 3 dB per doubling of distance [3].

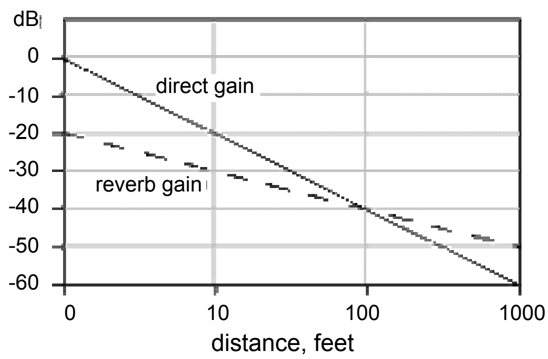

Figure 1: *An example of direct to reverb energy ratio [3].*

Many other researches also reported that reverberation is very important cue for distance, but the actual parameters for controlling perceived distance with reverberation cues are not well defined. But according to our previous research, we found that EDT (Early Decay Time) and C80 are varying systematically with distance [4], and these variations are due to early reflection decay curves.

In this paper, we will show a method of handling and shaping of early reflection decay curves for control of auditory distance, and psychological test results of auditory distance control with early refection decay curve shaping. With these validations, we can confirm early reflection decay curve shaping is effective for control of perceptual auditory distance and spaciousness of the sound sources.

2. BACKGROUND

Referring to previous research [4], we found that EDT (Early Decay Time) and C80 are varying systematically with distance. Namely, as the distance between source and listener goes farther, C80 more decreased and EDT more increased. Figure 2 and Figure 3 show the experiment results on this phenomenon. On the other hand, reverberation time of the specific room doesn't vary so much even if the measuring distances from sound source are increased.

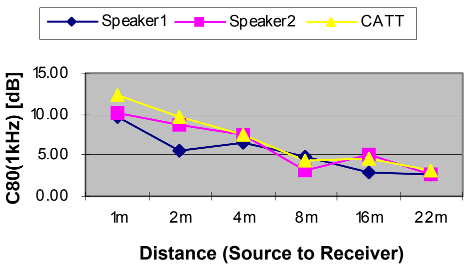

Figure 2: *C80 versus auditory distance.*

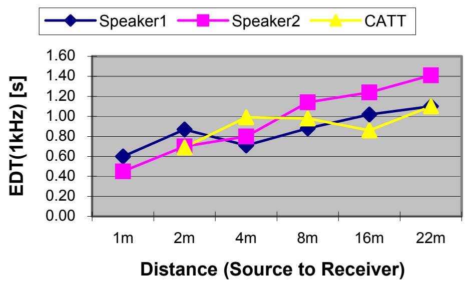

Figure 3: *EDT versus auditory distance.*

To investigate the reason that EDT and C80 are varying systematically with distance, we divided measured impulse responses into early reflection part and late reverberation part. And we normalized early reflection parts in order to compare their decay tendencies of each impulse response.

Figure 4 shows the decay curves of early reflections and rate reverberations for 1m and 16m distance. We can identify the decay curve slopes of early reflections are varying significantly, while the decay curves of rate reverberations have almost the same shape. Figure 5 shows the systematic changes of early reflection decay curve with distances.

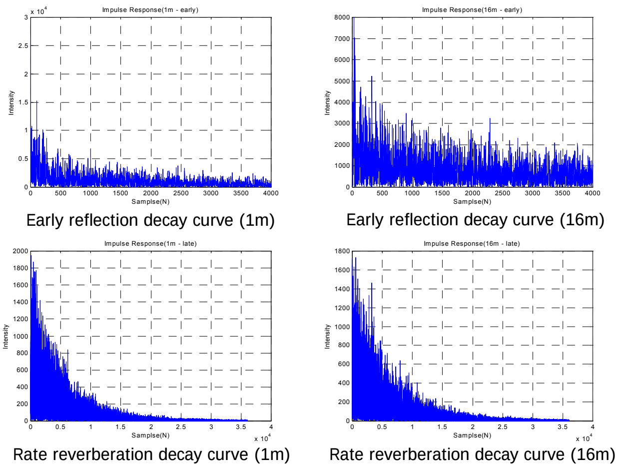

Figure 4: *Comparison of decay curve shapes.*

C80 is the energy ratio of early part and late part and EDT is the reverberation time for the first 10 dB drop. So these two parameters are very sensitive to the shape of the early decay curve. As the distance varies, the early reflection part varies significantly and this change of early reflection makes C80 decrease as the distance increases and vice versa in EDT. So if we can control these two parameters by designing the early parts of impulse responses, we can control the perceived distance with the reverberation cue. So C80 and EDT can be used as the parameters for controlling perceived distance with the reverberation cue.

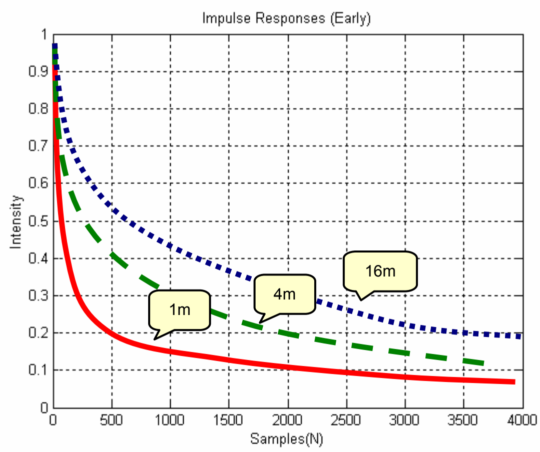

Figure 5: *Shapes of early reflections with distances.*

3. EARLY DECAY CURVE SHAPING

With previous results (systematic changes of EDT and C80 with distance) and specific measured impulse response, we tried to make simple function to provide distance information about a sound source for listening test. In analyzing measured impulse responses of a specific room, we found that late reverberation parts are maintained constantly whereas early reverberation parts vary systematically with distance. The slope of the early reflection curves of impulse responses measured at the near place of the source is more abrupt than the one at the distant place of the source.

From this investigation, we devised a method to implement situations above mentioned. First of all, we selected a measured 2m-impulse response as a reference impulse response. Then we made 1m-impulse response and 4m-impulse response using this reference impulse response. And we chose first 80ms as the threshold of early reflection part of the impulse response. This method can be divided into two parts. In the first step, we make late reverberation part according to various distance of the source. Namely, we chose late reverberation part of the measured impulse response at 2m in a specific room and used it as the late reverberation part of any other point of the room impulse response. Because late reverberation part of the impulse response has regular pattern with distances as we could see in measured data. Then it makes no difference to insert 2m-late reverberation part to 1m or 4m-late reverberation.

The second step is to make the early reflection part according to the distance change. We adjusted the EDT values with distance by varying the slope of early decay curve. When we adjust EDT, reverberation time (RT) must be fixed. If we consider a straight line with linear slope to measured 2m-impulse response appropriately, we can get some impulse responses that we want. But when we apply a linear straight line to early parts of the impulse response, there can be overemphasized reverberation and reverberation time of the room will not be fixed. Therefore we applied several linear lines of different slope for approximately shaping the curves of different distances. The curve can be represented by a pair of time and power value for the calculation of curves for the early reflection parts:

$$\{(t0, e0), (t1, e1), (t2, e2), \ldots, (tn, en)\}$$

Theoretically element t0 is the propagation time for the path from sound source to receiver. And elements e0, en can have the values of 2, 0 respectively. Then actual modification value for t0 (direct sound) at half distance will be 4 (square of 2) and that for tn will be 1 (0 power of 2).

When the position of sound source is changed by user interaction, the propagation time also varied. Thus we have to compensate this changed propagation delay, as insert or delete first samples of impulse response. The propagation time difference Δt for changing the position of sound source can be calculated with following numerical formula:

$$\Delta t = (d - ref\_d) / S,$$

where d is changed distance, ref_d is reference distance, and S is the velocity of sound propagation.

Figure 6 shows the way to control the early decay curve. These modified impulse responses are convolved with dry sound sources. The outcomes after convolving with distance are normalized to investigate the only effect of EDT excluding the influence of loudness. We have focused on how important the EDT value is in affecting the auditory distance with other acoustic parameters fixed like reverberation time. By means of outcomes above, we performed subjective listening tests with well-trained trainers to verify our hypothesis, that is, not the late parts but the change of early decay curve affect perceiving the distance of the source.

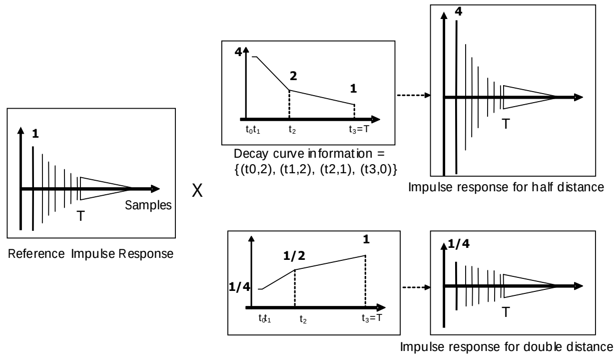

Figure 6: *Adjustment of early decay curve*

4. VALIDATION TEST

We chose dry sound sources of tenor solo, soprano solo and orchestra as test materials. Two kinds of subjective listening tests were performed. First, randomized the order of the different impulse responses and paired with two samples. Then let the subjects listen to the pair of samples, which are the same sounds with different distances, and asked them to choose which sample is perceived closer. Secondly, we arranged impulse responses in a distance order, e.g. 2m-4m-8m or 8m-4m-2m pair. Then we also asked the subjects whether they feel the sample becomes close or distant. In the first tests, when a subject finds a correct answer he gets +1 point, otherwise he gets –1 point. And in the second method, if subject finds a correct answer he gets +1 point and when he can't recognize any change, he gets 0 point. And if he answers incorrectly, he gets –1 point. So a perfect score is +3, and the lowest point is –3. Total 9 subjects participated in the subjective listening tests.

Figure 7 and Figure 8 show the results of subjective listening tests for two test methods respectively.

These graphs also show the total averaged scores of subjective responsiveness regardless of any kind of recording dry sound source. We tried to look into listeners' ability to perceive auditory depth on the average. So method 1 has a perfect score as +9 and the worst mark as –9, method 2 has a perfect score as +6 and the worst score as –6. Generally listeners had a higher mark in method 2. It shows that people perceive well in case of listening samples along to the distance in serial. Overall we obtained the satisfactory result through subjective listening tests in validating previous results, which explains how EDT affects the distance of the source.

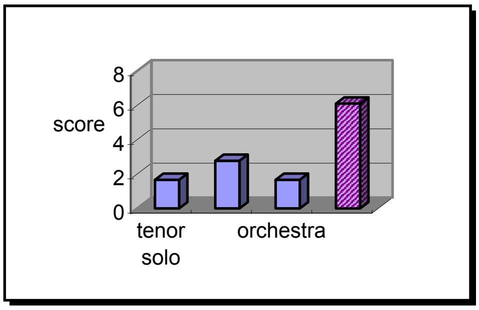

Figure 7: *Psychological test result according to dry sound sources (method 1).*

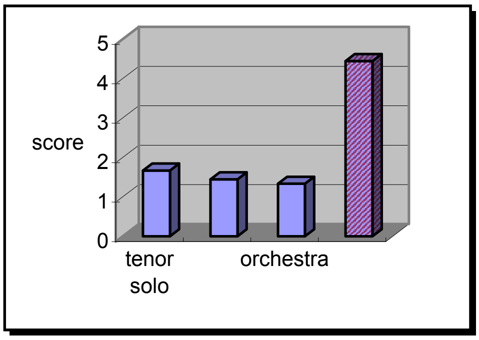

Figure 8: *Psychological test result according to dry sound sources (method 2).*

5. CONCLUSIONS

Many previous researches show that reverberation can provide absolute information. But the difficulty of implementation and no effectively defined parameters have prevented the reverberation from being used efficiently. We found EDT and C80 are varying systematically with the distance from sound source to receiver. This systematic change can be explained by examining the divided impulse responses. Because the early reflection part of impulse response varies significantly with distance whereas the late reflection part does not vary much, C80 that is the energy ratio of the early part over the late part of the impulse response decreases as the distance gets farther and vice versa in case of EDT. Both parameters are related to the reverberation of the room. So EDT and C80 can be used as the parameters for controlling perceived distance with the reverberation cue.

With this result and relatively simple method, we functionalized reverberation cue with distance. We made artificial reverberation curve composed of early parts with various slopes and late parts with the one specific slope to imitate impulse responses of different distances. The significance of this paper is to extract effective distance cues and to define EDT or C80 as distance cue maintaining reverberation time constant. To provide distance information of sound source in virtual audio environment, we need some practical algorithm. This paper also shows a good possibility to implement the reverberation parameter with proposed reverberation decay curve, which is composed of the slope of the early part changed and late part unchanged as distance variation.

For future work, we are going to investigate multi-channel impulse responses for multi-channel reproduction environment. And also we are going to make clear the time domain variation of early reflections of impulse response according to various positions of sound source in the room.

6. ACKNOWLEDGEMENTS

This work was supported by the Ministry of Information and Communication (MIC) of Korean government under the title of “Development of Super-intelligent Multimedia Anytime-anywhere Realistic TV (SmarTV) Technology”. We, authors, faithfully thank to the project funding managers.

7. REFERENCES

[1] Durand R. Begault, *3-D sound for virtual reality and multimedia*. Boston: AP Professional, 1994.

[2] Jens Blauert, *Spatial Hearing*, The MIT Press, 1983.

[3] William G. Gardner, “3D Audio and Acoustic Environment Modeling,” *Electronic Music and Audio Guide* at `www.sonicspot.com`, 1999.

[4] Han-Gil Moon, “Study on the Distance Cues in Virtual Audio Environment,” *Proc. Wespac*, 2003.