

Listening test for evaluating the similarity between artificial reverberation algorithms and real reverberation

Huan Mi, PhD.

Audio Lab, Department of Physic and Electronic, University of York

8th July, 2021

Purpose:

Reverberation algorithms can simulate different reverberation effect, but the closer to the real reverberation is, the better the reverberation algorithm is. This listening test is to find one or several reverberation algorithms that can simulate real reverberation as plausible as possible.

Materials:

webMUSHRA for scoring system about reference audio, test audios and anchor audio.

Methods:

1. There will be three big groups of listening tests about reverberation effect of three rooms with different reverberation time (0.266s, 0.95s and 2.34s). Each big group of listening test includes four small groups of trials with four different audio samples (male speech, female singing, solo cello piece and drumbeat) separately. The presentation of tests with different reverb times will be randomised. The reasons why I chose these audio samples are that male speech is used almost all applications, female singing as music sound source is familiar with public, solo cello piece is used for low frequency source, and drumbeat is used for slow bit transient and short impulsive sound. In each trial, there will be seven different reverberation audios generated by seven different reverberation algorithms, an anchor audio and a real reverberation audio as test audios, and the real reverberation audio will be as the reference audio. These seven reverberation algorithms are generated by following reverberation algorithms separately:

- a) Schroeder reverberation algorithm
- b) Gardner reverberation algorithm
- c) Moor reverberation algorithm
- d) Feedback delay networks
- e) Directional feedback delay networks
- f) Dattorro reverberation algorithm
- g) New reverberation algorithm designed by myself

2. In each trial, participants need to listen to the reference audio and test audios carefully, and score the test audios according to their perceptual similarity with the reference audio on a scale of zero to a hundred. The most important thing is that they need to compare which test audio is same as the reference audio and score a hundred, and the more similar the other condition audio samples are to the reference audio sample, the higher the score for them.

3. About 20 participants listen to the reference audio and test audios respectively and rate the test audios according to their perceptual similarity with the reference audio on a scale of zero to a hundred.

4. The experimental results will be summarized and analysed.

Data Interpretation:

Four histograms will be used to plot the results. The seven reverberation algorithms will be plotted on the X axis (abscissa) and each algorithm refers to three different reverberation time (0.31s, 0.91s and 1.51s) respectively. The average score will be on the Y axis (ordinate). Error bar with standard error will be plotted. Then the score of these seven reverberation algorithms will be analysed with ANOVA test in different reverberation time,

different reverberation algorithms and different audio samples to decide whether they have significant difference. If the average score of one of these seven algorithms higher than others and their results have significant difference, we can conclude that the perceptual reverberation generated by the reverberation algorithm is the most similar with the real reverberation. Besides, the age and gender of the participants will be collected.