Audio Engineering Society

Conference Paper

Presented at the Conference on
Spatial Reproduction
2018 August 6 – 9, Tokyo, Japan

Comparison of Microphone Distances for Real-Time Reverberation Enhancement System Using Optimal Source Distribution Technique

MARUI, Atsushi[1], YAIRI, Motoki[2], and KAMEKAWA, Toru[1]

[1]*Faculty of Music, Tokyo University of the Arts*
[2]*Kajima Technical Research Institute*

Correspondence should be addressed to MARUI, Atsushi (marui@ms.geidai.ac.jp)

ABSTRACT

Although musical instrument players may want to learn the spatial impressions of the concert hall beforehand, practicing in the actual hall is not practical due to the geographical and the financial reasons. Hence, it is helpful for players if there is a spatial reproduction system that can recreate the concert hall reverberation in their practice environments. An implementation of crosstalk cancellation method for synthesizing virtual auditory space using loudspeakers called Optimal Source Distribution technique is used to add realistic reverberation of a live venue to a violin performance in real-time. In this research, one of the defining parameters for the timbre of convolved reverberation, microphone placements on and over the instrument, were compared by professional players.

1 Background

When practicing a musical instrument towards a live performance in a concert hall, the player may want to learn how the spatial impressions of the room affect the performance before the actual concert. However, in many cases, practicing in the actual concert hall is not practical due to geographical and financial reasons. Hence, it is helpful for such musical instrument players if there is a spatial reproduction system that can recreate the concert hall reverberation in their practice environment.

Such attempts were made by supplying artificial reverberation to a small recital hall to simulate a larger concert hall (*e.g.*, [1]) or convolving impulse responses of a concert hall and playback the reverberation from loudspeakers in an anechoic chamber or a practice room (*e.g.*, [2, 3] and [4], respectively). These methods employ a number of loudspeakers that have to be spatially distributed in a practice environment, which may not be the optimal solution for casual users who do not wish to install a large system in their practice rooms. For that reason, supplying binaurally processed reverberation through a pair of headphones may become a solution at the other end of the spectrum of playback systems. However, headphone reproduction of binaural reverberation has an obvious drawback that the experience of wearing headphones is different from the actual concert situations, and furthermore, headphone enclosures may physically interfere with the certain family of instruments such as violins.

Optimal Source Distribution (OSD) technique is an implementation of crosstalk cancellation method for synthesizing virtual auditory space using horizontally placed loudspeakers [5]. Using such an implementation, real-time 3D sound reproduction system for binaural synthesis over loudspeakers can be achieved. As a part of our ongoing research, an OSD implementation was used to add realistic reverberation of a live venue to a violin performance in real-time to achieve the goal stated above.

In this context, one of the defining parameters for the timbre of convolved reverberation is in the microphone setup on the instrument. In this research, several microphone placements on and over the instrument were compared by professional violinists.

2 Method

2.1 Impulse Response Measurement

In order to provide an instrument player with the hall reverberation using a binaural reproduction system, a signal processing chain that contains both the room-related impulse response and the head-related impulse response of the instrument and the player is necessary. For that reason, the binaural room impulse response was measured prior to the evaluation experiment.

The measurement was done in Studio A of the Tokyo University of the Arts (volume: 1060 m^3, surface area: 700 m^2, floor area: 170 m^2, reverberation time at 500 Hz: 0.97 seconds). Fig. 1 shows the placement of the loudspeaker (Eclipse TD508) and the dummy head microphones (Neumann KU100) that follows the player and the instrument positions used in a concert situation of a violinist.

Recording was done using Pink TSP signal [6] of 2^{17} samples long in 48 kHz sampling frequency and 24 bit resolution. Pink TSP is a variant of measurement methods using a logarithmically-swept sine wave signal.

The impulse response was subsequently processed so that the data can be used with the OSD system. Specifically, deleting the direct sound and delay compensation for the experimental setup were done.

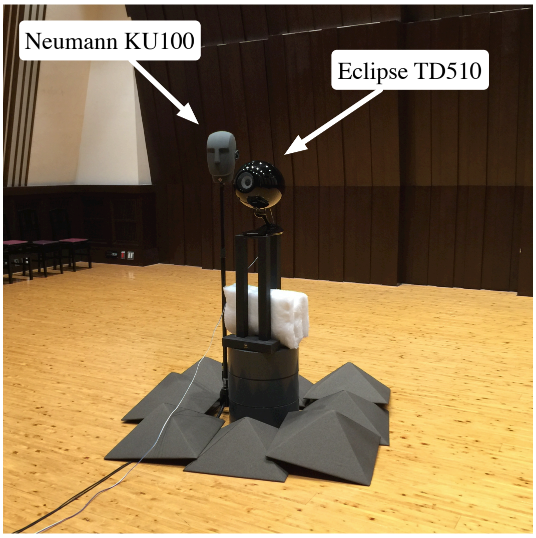

Fig. 1: Placements of the loudspeaker and the dummy head in Studio A of the Tokyo University of the Arts.

2.2 Evaluation Experiment Setup

The room used for capturing the head-related impulse response (Studio A of the Tokyo University of the Arts) and the adjacent control room was used for the experiment. OSD loudspeakers were placed inside the control room with additional sound absorbers to reduce reflections from the wall surfaces. Fig. 2 shows the OSD speaker and the participant positions. The head of the violin player was set to 1.5 meters facing towards the center of the loudspeakers, and position and posture of the instrument and the body were set accordingly.

One microphone (DPA 4099, super-cardioid) was attached near the sound hall of a participant's violin using a DPA VC4099 clip. Three microphones were set to have distances of 0.15, 0.30, and 0.75 meters from the bridge of the instrument (Schoeps CMC64 (cardioid) for 0.15 and Schoeps CMC641 (super-cardioid) for 0.30 and 0.75 meters positions). All microphones were placed on normal line on the violin surface and pointed towards the bridge. Since each violinist held the violin in different orientation, microphone placements were adjusted accordingly. (Fig. 3)

The microphones were connected to a four-channel microphone amplifier, which is then connected to a digital computer equipped with a signal processing

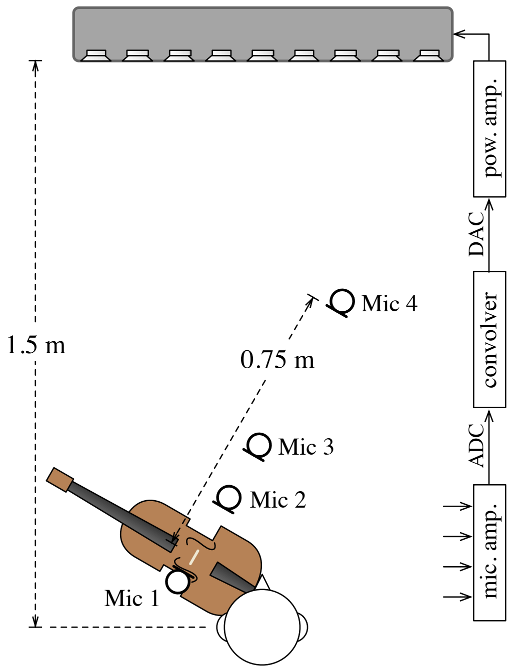

Fig. 2: Experimental setup. A violinist, nine-way OSD speaker, and other equipment are shown. Cables from the microphones to the microphone amplifier are omitted.

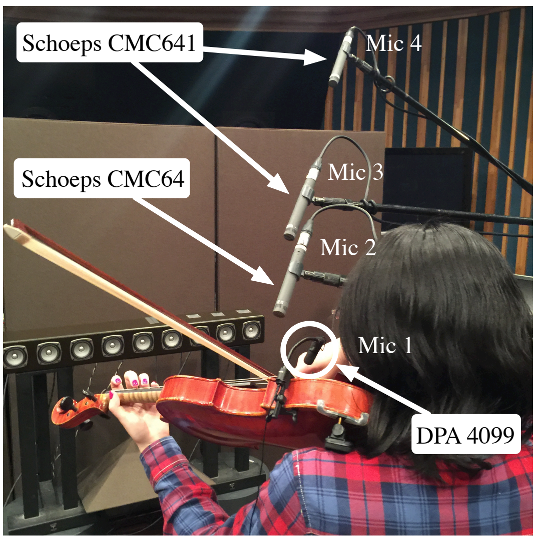

Fig. 3: Photograph of microphone placements in the experiment. All microphones were placed on normal line on the violin surface and pointed towards the bridge.

software for impulse response convolution. The input violin signal was convolved with the binaural impulse responses measured in the studio, and sent to the power amplifier connected to the OSD loudspeakers. No delay compensation was done for the different microphone distances.

The power amplifier was adjusted to recreate the level of the reverberation at the player's ears. Note that the violinists heard the direct sound of the actual violin and all other reflections were binaurally provided from the OSD loudspeakers, if the space was anechoic. However, although the experimental setup was done in the fairly dry room, the output levels from the OSD speakers had to be lowered by 15 dB to avoid the feedback into the farthest placed microphone.

2.3 Task

The task of the participants was to compare the two among five levels of stimuli (including one actual live room and four microphones used with OSD reproduced environments) and respond of which and by how much one stimulus was rated higher over the other stimulus in six 11-point attribute scales. All possible pairs of the five stimuli were compared against each other.

The attributes used were "timbre (*onshoku*)," "spatial impression (*kuukan inshou*)," "reverberation (*zankyou-kan*)," "ease of playing (*ensou no shiyasusa*)," "overall quality (*sougouteki-na yosa*)," and "preference (*kojin-teki-na konomi*)," with corresponding Japanese attributes in parentheses. The Japanese attributes (shown in *italic*) were used in the experiment, because all violinists spoke native Japanese.

Because the test method used was the pairwise comparison among five stimuli, there were originally 20 comparisons ($= 5 \times (5-1)$) that required a total of 40 evaluations of the stimuli. That was sought to result in excessive fatigue to complete the task. In order to reduce the number of pairwise comparisons, Eulerian circuit [7] was employed. That is, a stimulus was always evaluated compared to the stimulus which the participant was just previously evaluating. For example, when a participant is evaluating A, B, and C in this order, the participant was asked to evaluate B in comparison to A and then evaluate C in comparison to B (as shown in the trial numbers 1 and 2 in Fig 4). Fig 4 only show the first 10 comparisons, but the second half of the comparisons were made

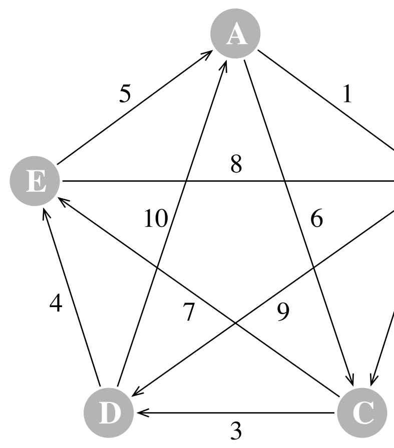

Fig. 4: Stimulus presentation order. Five stimuli were assigned to five circles labeled with alphabets (A to E). Arabic numbers (1 to 10) on arrows show the trial numbers to compare the two corresponding environments denoted by the attached circles.

by following the paths in reverse order (which makes A-B-C-D-E-A-C-E-B-D-A-D-B-E-C-A-E-D-C-B-A evaluation order). This reduced the number of evaluations to 21 instead of 40 while keeping the necessary 20 comparison pairs. The five stimuli were semi-randomly assigned to A through E with an idea from balanced Latin square designs.

2.4 Participants

Three professional violinists from Tokyo University of the Arts Philharmonia Symphony Orchestra participated in the experiment. All violinists have more than 25 years of violin experience. The instruments used in the experiment were the violins owned by each of the participants. All violinists were compensated for participating in the experiment.

3 Results

The analysis was done following the Ura's variation of Scheffé's ANOVA on Paired Comparison [8, Ch. 19]. Average ratings on five stimuli for six attributes are shown in Fig. 5. "A" stands for Studio A and each of "1" through "4" stand for the microphone from closely placed to farther placed (corresponds to the ones shown in Fig. 2 and Fig. 3). The thin lines connecting symbols show statistically significant differences between the two average ratings with $\alpha = .05$ level. Statistically significant differences between the live room and the simulated environment (four microphones) were found in many of the attributes. The live room ("A") was always rated higher than the other simulated environment.

Among the four microphone placements ("1"–"4"), the farther placed microphones were almost always rated higher in all attributes. This can be explained from at least two viewpoints. First, the microphone farther from the instrument can capture wider area of sound radiated from the instrument, resulting in more "natural" sounding reverberation for the instrument players. Second, although the room we used for the experiment was treated to reduce reflections off the walls, it was not an anechoic environment. The early reflections of the room indeed were captured by the microphones and resulted in providing more reverberation to the violinists.

The violin mounted microphone (shown as "1" in Fig. 5) was rated the lowest in most attributes except for "spatial impression" (but with no significant differences within the four microphones). The microphone had statistically significant differences against the other microphones in attributes "timbre," "overall quality," and "preference." However, the differences seemed to arise from the microphone placement distance rather than the brand or the model of the microphone.

It was our surprise to find that there was no statistically significant differences between "A" (in the studio) and "4" (the farthest microphone) in the attributes "reverberation," "ease of playing," "overall quality," and "preference." Although not being able to reject the null-hypothesis is likely to be due to the small number of participants, it still is encouraging for the reverberation enhancement system that the score differences between the actual room and the top-rated microphone are within the comparable range (*i.e.*, score differences between top-rated and bottom-rated microphones are larger in many of the attributes).

4 Summary

Because practicing in the actual concert hall is not practical due to geographical and financial reasons, it is helpful for such musical instrument players if there is a spatial reproduction system that can recreate the concert hall reverberation in their practice environment.

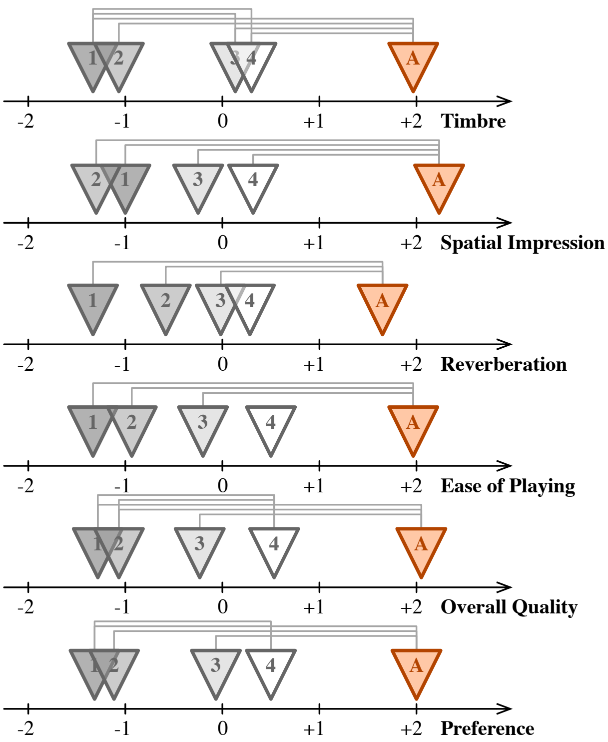

Fig. 5: Average ratings on five stimuli for six attributes. "A" stands for Studio A and each of "1" through "4" stand for the microphone from closely placed to farther placed (corresponds to the ones shown in Fig. 2 and Fig. 3). Thin lines connecting symbols denote there was a statistically significant differences between the two average ratings.

Using an Optimal Source Distribution (OSD) implementation which is a crosstalk cancellation method for synthesizing virtual auditory space using horizontally placed loudspeakers, real-time 3D sound reproduction system for binaural synthesis over loudspeakers can be achieved. Especially, an OSD implementation can also be used to add realistic reverberation of a live venue to an instrument performance in real-time. In this research, one of the defining parameters for the timbre of convolved reverberation, microphone placements on and over the instrument, were compared by professional violin players. Also, the actual venue was included in the comparisons.

It was a rather obvious finding that the actual live room was rated highest with statistically significant differences against the simulated environment. However, the score differences between the live room and the microphone placements were within the comparable range, notably for the attributes "ease of playing," "overall quality," and "preference." Placing the microphone farther away from the instrument decreases the direct-to-reverberation ratio at the playing/listening position, resulting in increasing overall ratings for all attributes related to timbre, space, and preference, but with a higher risk of feedback loops within the reverberation enhancement system. The result from this study may provide a good trade-off between microphone distance and timbral/reverberation naturalness in the reverberation enhancement system such as the one used in this research.

References

[1] Lokki, T., Patynen, J., Peltonen, T., and Salmensaari, O., "A Rehearsal Hall with Virtual Acoustics for Symphony Orchestras," in *Proceedings of 126th AES Convention*, Audio Engineering Society, 2009.

[2] Ueno, K., Sato, F., Tachibana, H., Ono, T., and Senju, M., "Subjective Experiment on the Effect of Hall Response on Violin Player in Simulated Sound Field," *Journal of Acoustical Society of America*, 100, p. 2837, 1996.

[3] Ueno, K. and Tachibana, H., "Experimental Study on the Evaluation of Stage Acoustics by Musicians Using a 6-channel Sound Simulation System.

Acoustical Science and Technology," *Acoustical Science and Technology*, 24(3), pp. 130–138, 2003.

[4] Tsuchikura, R., Nakahara, M., Mikami, T., Kamekawa, T., and Marui, A., "Acoustic Properties of Practice Rooms Where Musicians Can Practice Contentedly: Effect of Reverberation on Practice," in *Engineering Briefs of 133rd Convention*, Audio Engineering Society, San Francisco, United States, 2012.

[5] Takeuchi, T. and Nelson, P. A., "Optimal source distribution for binaural synthesis over loudspeakers," *Journal of Acoustical Society of America*, 112(6), pp. 2786–2797, 2002, https://doi.org/10.1121/1.1513363.

[6] Suzuki, Y., Asano, F., Kim, H.-Y., and Sone, T., "An Optimum Computer-Generated Pulse Signal Suitable for the Measurement of Very Long Impulse Responses," *Journal of Acoustical Society of America*, 97(2), pp. 1119–1123, 1995.

[7] Hagita, M., "Elegant na Kaitou wo Motomu (A Request for Elegant Solutions, Answer to the Problem from April Issue)," *Suugaku Seminar (Mathematics Seminar Magazine)*, 54(7), pp. 88–93, 2015, contents are written in Japanese language.

[8] Satou, M., *Toukeiteki Kannou Kensahou (Statistical Methods for Sensory Tests)*, Nikka Giren (Union of Japanese Scientists and Engineers Press), 1985, ISBN 4817190116, contents are written in Japanese language.