ISRA 2019 AMSTERDAM

PROCEEDINGS of the
International Symposium on Room Acoustics

15 to 17 September 2019 in Amsterdam, Netherlands

A study on precise measurement of room impulse response in a scale model and auralization

Kazunori SUZUKI and Takayuki HIDAKA[1]

Takenaka R&D Institute, Japan

ABSTRACT

As a sequence of the authors' previous work [1, 2], this paper discusses fundamental issues of the scale model experiment and reformulates the recording procedure of the ambisonics in a scale model. Basic psychological test on sound localization is reported to verify its applicability.

Keywords: Scale model experiment, Auralization, Physical phenomena

1. INTRODUCTOIN

At a room impulse response (RIR) measurement in a 1/10 scale model, ten times the temporal resolution of the full size is required under the limitation of hardware such as a microphone and a sound source. When auralization is intended, the RIR has to fulfill more precise conditions in terms of the audio frequency range, the duration, and the dynamic range than the objective parameter measurement. Accordingly, further pre and/or post processing are required beyond the conventional measurement procedures. In this report, focusing on the air absorption, a trade-off of the octave band analysis is described. Next a corrective method is proposed which enables to express the RIR as a function of continuous frequency. Based on this, the ambisonics recording technique is revised to measure the RIRs suitable for auralization in a scale model.

2. FUNDAMENTAL ISSUES

2.1 Air absorption

Take the reverberation time for instance. As the excessive absorption by air molecules at higher frequencies results in a substantial error on its accuracy, the dry air or nitrogen substitution method [3] has been used often in a 1/8 to 1/12 scale model. However, such a method is costly or involves a dangerous procedure in practice. Another alternative is numerical correction to the measurement data taken under the normal air condition, avoiding the medium control.

Then, the envelope of the acoustic energy decay process in a model is approximated by an exponential function,

$$E_m(t) = E_m(0)\,\mathrm{e}^{-6\ln 10 t/RT_m}\ , \qquad (1)$$

and the corrected energy decay is written as,

$$E_c(t) = E_m(t)\,\mathrm{e}^{(m_m - m_c)ct} = E_m(0)\,\mathrm{e}^{(-6\ln 10/RT_m)t}\ \mathrm{e}^{\Delta mct}\ , \quad \Delta m = m_m - m_c\ . \qquad (2)$$

Here, suffixes m and c mean the measurement in a model and the compensated, m_m is the attenuation constant of air at the measurement site and m_c is usually taken the value of 20°C and relative humidity 2%. RT is a reverberation time. Accordingly, the following simple relation holds.

$$RT_c = \frac{RT_m}{1 - \Delta m(c/6\ln 10)RT_m} \qquad (3)$$

It is seen that the reverberation time ought to be 1.5 sec at 5 kHz in the model when one aims RT = 2 sec at 500 Hz in an actual space (Fig. 1). The difference between these RTs is about 35%. At 20 kHz, the measurement range of 0.3 ~ 0.4 sec in a model expands to 1~3 sec in the actual hall. Taking into account of the ambiguity in the decay estimation to the second decimal place (even if Schroeder integration is applied), this result indicates RT measurement at higher frequencies is unmanageable in practice. Just in case, there are observable upper limiting frequencies of the RT_m, which is defined as that sets zero to the denominator

[1] hidaka.takayuki@takenaka.co.jp

of Eq. (3). They are 0.46 and 0.16 sec at 20 and 40 kHz, respectively.

As a matter of course, the same goes for energetic parameters such as C80 and G. One idea to solve this problem to some extent is discussed in the later chapter.

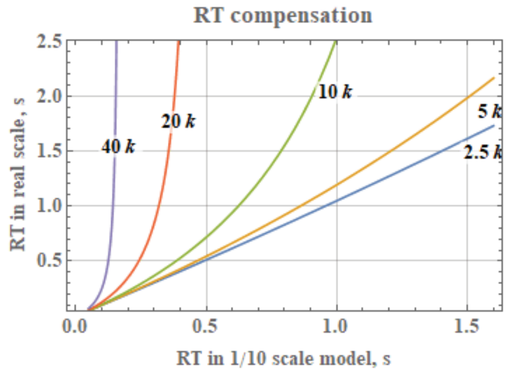

Fig. 1 Effect on RT by the numerical compensation of dry air method.

2.2 Acoustic boundary layer

The interior walls in a hall can be assumed rigid excepting seating areas and an intentional absorptive area for sound control. In a scale model, the acoustic boundary layer above a solid surface has an apparent admittance [4] which gives systematic error in the predicted sound field,

$$\beta = \frac{1-i}{2^{1/2}}\sqrt{\frac{\omega\mu}{\rho c^2}}\left(\sin^2\theta + \frac{\gamma-1}{\mathrm{Pr}^{1/2}}\right) \tag{4}$$

Pr is the Prandtl number, γ is the ratio of the specific heats, ρ is the air density and μ is the viscosity of air. For air in the temperature range 0 to 40°C, $(\gamma-1)/Pr^{1/2}$ is 0·48. Accordingly, the random incident absorption coefficient of the boundary layer is calculated as,

$$\alpha_{stat} = \int_0^{\pi/2}\left(1-\left|\frac{\cos\theta-\beta}{\cos\theta+\beta}\right|^2\right)\sin 2\theta d\theta\,. \tag{5}$$

Figure 2 is a plot of Eq. (5). In a scale model experiment, this term should be considered separately from the residual absorption [5] in the reverberation time equation, that is,

$$RT_m = 0.161\frac{V}{S\bar{\alpha}+4m_m V + S_{res}\alpha_{stat}}\,. \tag{6}$$

Figure 3 shows a ratio of the reverberation times between without and with the boundary layer term in Eq. (6). Here, assuming a shoebox hall with 1,600 seats, we set the room volume V = 16,000m^3 and residual absorption area S_{res} = 5,000m^2. From this, when the RT observed in a model is 1.5 sec, the resultant RT has about 10% and 20% error respectively at low and mid frequencies if the boundary layer term is dropped. When assembling a scale model, we have to evaluate the total absorption area beforehand. We should take a precaution about the error caused by the boundary layer.

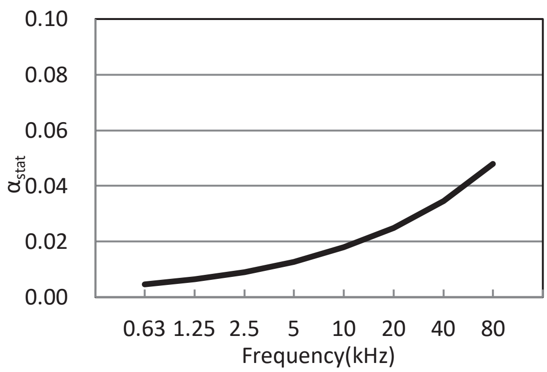

Fig.2 Statistical absorption coefficient of the boundary layer.

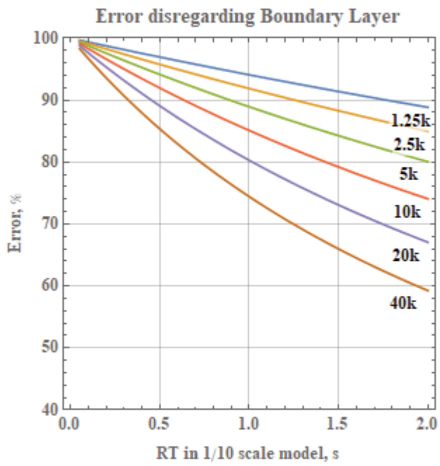

Fig. 3 Error in RT estimation when the boundary layer dissipation is dropped.

3. SIGNAL PRCESSING FOR AURALIZATION

Experimental apparatus and environment can degrade the measurement accuracy in a scale model more often than in actual hall. Some post signal processing that corrects the RIRs favorable for auralization is proposed.

3.1 Dissipation by air

The amplitude attenuation of sound wave against propagation distance caused by air molecules is written with exponential function, so that one can formulate this phenomenon in the following introducing the Fourier transform pairs:

$$\phi(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{dk(\omega)}{d\omega} S(\omega)\ \mathrm{e}^{ik(\omega)x} d\omega \tag{7}$$

$$S(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(x)\ \mathrm{e}^{-ik(\omega)x} dx \tag{8}$$

In Eq. (7), $S(\omega)$ is the Fourier transform of RIR measured in a scale model, where $k = \omega/c + i\alpha$, ω, c, α are the wave number, the angular frequency, the sound velocity and the attenuation coefficient by air molecules, respectively. Taking into account of $\exp(-ikx)$ being the Green function in one dimensional sound field, $\phi(x)$ can be interpreted as a spatial impulse response for a virtual sound source located at x, and is calculated by Eq. (7) substituting α_m value at the measurement air condition. In other words, $\phi(x)$ represents a continuous amplitude distribution function along one dimensional axis. Again, substituting $k = \omega/c + i\alpha_c$ value at desirable temperature and humidity condition into Eq. (8), the spectral representation of RIR $S_c(\omega)$ is given. Finally, the RIR in time domain is obtained by taking the Inverse Fourier transform to $S_c(\omega)$. For 1/10 model, we can introduce a new variable $x' = 10x$ and α value in the actual space into Eq. (8).

At the actual calculation of Eq. (7), following equation should be considered.

$$\frac{dk(\omega)}{d\omega} = \frac{1}{c(\omega)} + \omega\frac{d}{d\omega}\left[\frac{1}{c(\omega)}\right] + i\frac{d\alpha(\omega)}{d\omega} = \frac{1}{c(\omega)} - \frac{\omega}{c^2(\omega)}\frac{dc(\omega)}{d\omega} + i\frac{d\alpha(\omega)}{d\omega} \tag{9}$$

For numerical purpose, next approximation is reliable where attenuation α is given by [6].

$$\frac{dk(\omega)}{d\omega} \cong \frac{1}{c} + i\frac{d\alpha(\omega)}{d\omega} \tag{10}$$

3.2 Amplitude frequency dependence

When dodecahedral loudspeakers are used as the sound source, its directional pattern strongly fluctuates above the cut-off frequency of the omnidirectionality [7] so that a certain compensation of the radiated sound is required. One practical way to realize a reasonable flat response all over the frequency range is to adjust its amplitude response the same as that by a high-fidelity one-way loudspeaker in a reverberant chamber. We used a parametric equalizer for digital editing for this purpose. In actual application, it is recommended this equalizer response is convolved with the anechoic music in advance.

3.3 Dynamic range

Auralization requires wider dynamic range than usual objective parameter measurements, say 60dB or more. However, such value is too wide to attain at usual RIR measurements and some sort of post signal processing is necessary. If later part of the RIR is influenced by back ground noise of the measurement system, this component extends the subjective reverberation than the true value after convoluted with music signal, especially at the stop cord of music.

For this purpose, we decomposed the measured RIRs into 1/3 octave bands and expanded the dynamic range at each band by linear extrapolation as shown in Fig. 5. Here we employed the average slope of early part from -5 to -25 dB range. This procedure is based on the truth that the reverberation component has the Gaussian noise statistics [8, 9] and phase information of the later reflections does not substantially influence on listener's subjective impression [10].

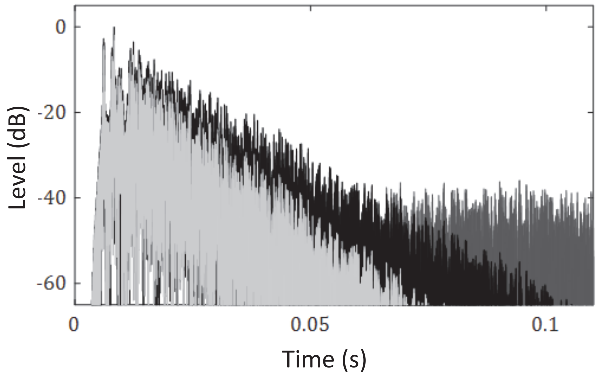

Fig. 5 Expansion of the dynamic range of RIR. Dark gray: original decay; light gray: extrapolated in accordance with the early part; black: the air molecules and boundary layer losses are compensated.

4. AMBISONICS

4.1 Reformulation

We assume a plane wave of amplitude Q incidents from the azimuth angle ψ and the elevation angle ϕ (with overhead as 0°). Then the sound pressure at a receiving position $\mathbf{r}(r,\theta,\varphi)$ is expressed by the spherical harmonics series,

$$p(\mathbf{r},k) = 4\pi Q \sum_{n=0}^{\infty} X_n(k\mathbf{r}) \sum_{m=-n}^{n} Y_n^m(\theta,\varphi) \cdot Y_n^m(\psi,\phi)^* \tag{11}$$

For U-channel receiving signal $\mathbf{p} = [p(\mathbf{r}_1,k),\cdots p(\mathbf{r}_U,k))]^t$, this equation is written in the matrix form with truncated order N:

$$\mathbf{p} = 4\pi Q \mathbf{X}\mathbf{Y}_r \mathbf{Y}_s^* \tag{12}$$

where $X_n(k\mathbf{r}) = i^n j_n(k\mathbf{r})$, $j_n(k\mathbf{r})$ is a spherical Bessel function, and Y_n^m is a spherical harmonic function. The detailed matrix definition is given by eq. (5) of [2].

When the receiving position $\mathbf{r}$ are chosen so as to make each row in the matrix $\mathbf{XY}_r$ mathematically independent, the ambisonics B signal is obtained by multiplying the pseudo inverse matrix $(\mathbf{XY}_r)^\dagger$ from the left side,

$$4\pi Q \mathbf{Y}_s^* = (\mathbf{XY}_r)^\dagger \mathbf{p} = \mathbf{B}, \quad \mathbf{B} = [B_0^0, B_1^{-1}, B_1^0, B_1^1, \cdots]^t \tag{13}$$

For $N = 1$, the components of **B** are commonly expressed by W, X, Z, Y.

At the decoding stage, the reproduced sound pressure is also expressed by the spherical harmonics, when L number of loudspeakers are laid out on a spherical surface,

$$p(\mathbf{r},k) = 4\pi \sum_{n=0}^{\infty} \sum_{m=-n}^{n} X_n(k\mathbf{r}) \sum_{l=1}^{L} a_l(k) Y_n^m(\theta_l,\varphi_l)^* Y_n^m(\theta,\varphi) \tag{14}$$

Here (θ_l, φ_l) is the direction of the l-th loudspeaker and a_l is the input signal to it. In a similar way, eq. (14) is expressed in the matrix form by the truncation [2].

$$\mathbf{p} = 4\pi \mathbf{a}(\mathbf{XY}_r)\mathbf{Y}_l^* \tag{15}$$

Finally equating Eq. (11) to Eq. (12), the input signal to the speaker $\mathbf{a} = [a_1(k), a_2(k), \cdots a_L(k))]^t$ is given.

$$\mathbf{a} = 1/(4\pi)(\mathbf{Y}_l^*)^\dagger (\mathbf{XY}_r)^\dagger \mathbf{p} = 1/(4\pi)(\mathbf{Y}_l^*)^\dagger \mathbf{B} \tag{16}$$

4.2 Microphone arrangement

To minimize errors at the RIR measurement, that is $p(r, k)$ of Eq. (11), a single 1/8 inch microphone (B&K, type 4138) is shifted sequentially on an automatic stage (Chuo Precision Industrial, ALD-4011-G0M + LV- 4042-1, https://www.chuo.co.jp/english/) in the sound field. The positioning accuracy of this automatic stage is within 0.03 mm, which is 1/100 times the wavelength of 100 kHz sound wave.

For the 1st order ambisonics (N = 1), seven microphone positions are chosen. This coincides with the ambisonics B-format sound recording. They are the target receiving point (No.1 in Fig. 6) plus six symmetrical points (No.2 to 7) along x, y and z axes around it. In order to obtain suitable S/N ratio, the total frequency range is subdivided into three to four. Therefore separate RIRs are measured by changing the distance d between microphones facing each other so as to avoid the spatial aliasing at each frequency range (Table I). For the second order (N = 2), a total of 11 positions are used including four additional points (No. 8 to 11) instead of the mathematical requisite number 9. By these modifications, the numerical errors in the pseudo inverse matrix estimation can be improved.

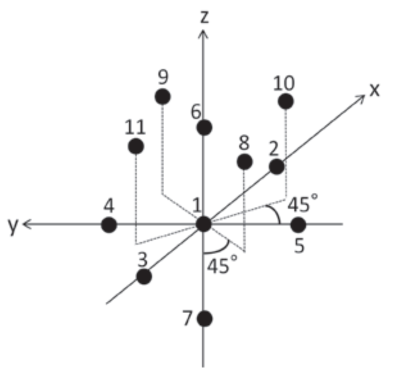

Fig. 6 Microphone positions in (x, y, z) coordinate. From no.1 to 11: (0,0,0), (d/2,0,0), (-d/2,0,0), (0, d/2,0), (0, -d/2,0), (0,0, d/2), (0,0,- d/2), (-d/a, -d/a, d/2), (d/a, d/a, d/2), (d/a, -d/a, d/2), (-d/a, d/a, d/2), $a=2^{3/2}$.

Table I Distance d between the microphones facing each other and the upper frequency within 1 dB error by the finite difference approximation.

Distance d (mm)	16	8	4	2	1
Upper frequency within 1dB error (kHz)	6	11	22	45	89
Corresponding freq. band in 1/10 model (kHz)	～6.3	5～12.5	5～25	20～50	20～100

4.3 Psychological test

Simple experiment was carried out to see if the 1/10 scale ambisonics had equivalent localization reproducibility to the real scale measurement. Impulse responses from a small loudspeaker were recorded by 15° step in horizontal and vertical planes in an anechooic chamber. Ambisonic signal were reproduced by six loudspeakers in horizontal and vertical plane, respectivelly. Pink noise (6 sec) and anechoic violin solo (cadenza from Mozart violin concert No. 4; 8 sec) were used as the test signal. Fourteen male and female subjects with normal hearing ability (aged from 29 to 49) joined the psychological test.

Figure 7 shows almost the same accuracy of localization for the scale model and full scale experiments. Similar results were obtained throughout this experiment for the first and second order ambisonics. Subjects expressed that there was no audible noise in the reproduced sound, which meant either the 1/10 or the full scale ambisonics had an equivalent sound quality. The error rate in the front-back judgment was slightly lower for the pink noise than for the violin music. This was supposed to be caused by the difference in the signal's spectra.

5. CONCLUSIONS

Signal processing beneficial to obtaining subjectively rendered RIR in a 1/10 scale model is described. The modified microphone arrangement for the ambisonics recording is proposed, which minimizes the experimental error in the RIRs. Psychological test with regard to the localization shows the scale model ambisonics up to the second order has identical performance with the real scale measurement.

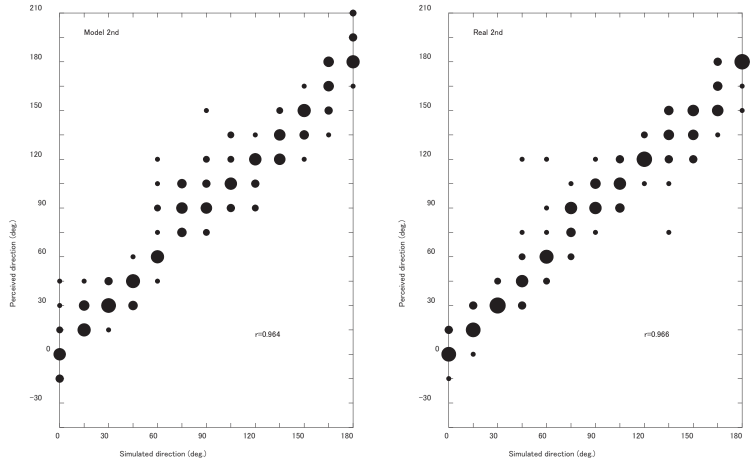

Fig. 7(1) Experimental result of the localization in the horizontal plane. Signal: pink noise; second order ambisonics. Left: 1/10 scale model, Right: Real scale.

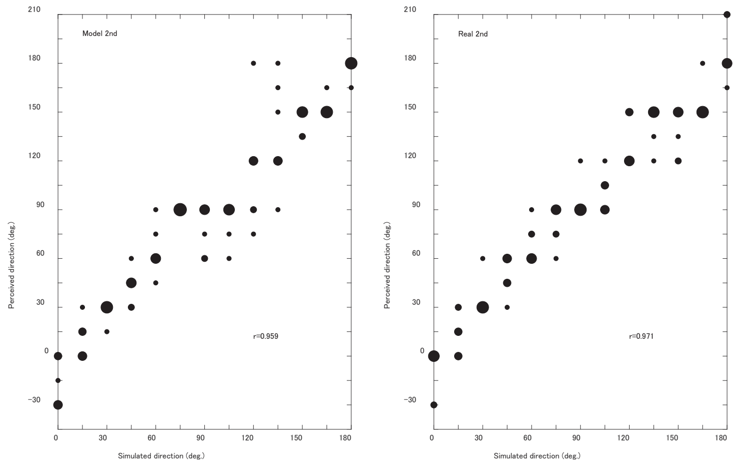

Fig. 7(2) Key same as Fig. 7(1), but signal: violin solo.

REFERENCES

1. T. Hidaka, et al. (2010), "A new miniature loudspeaker for room acoustical scale model experiment," Proc. 20th I.C.A., Sydney.
2. K. Suzuki, et al. (2018), "Basic study on improvement of precision of measurement of room acoustic characteristics using scale model and three-dimensional sound field auralization," J. A. S. Jpn. **74**, 244-253 (in Japanese).
3. T. Gade (2014), *Springer Handbook of Acoustics* (Springer Science + Business Media) chap. 9.5.5.
4. A. D. Pierce (1981), *Acoustics* (McGraw-Hill, New York), pp. 508–534.
5. L. L. Beranek and T. Hidaka (1998), "Sound absorption in concert halls by seats, occupied and unoccupied, and by the hall's interior surfaces," J. A. S. A., **104**, 3169-3177.
6. H. E. Bass et al. (1995), "Atmospheric absorption of sound: Further developments," J. A. S. A., **97**, 680-683
7. V. Tarnow (1974), "Sound radiation from loudspeaker system with the symmetry of the platonic solids," B&K Tech. Rev. **4**, 23-31
8. H. Kuttruff, "On the audibility of phase distortion in rooms and its significance for sound reproduction and digital simulation in room acoustics," Acustica, **74**, 3–7 (1991).
9. T. Hidaka, Y. Yamada and T. Nakagawa, "A new definition of boundary point between early reflections and late reverberation in room impulse responses," J. A. S. A., **122**, 326–332 (2007).
10. R. Heinz, "Binaural room simulation based on an image source model with addition of statistical methods to include the diffuse sound scattering of walls and to predict the reverberant tail," Appl. Acoust., **38**, 145–159 (1993).