5aPPa7

PERCEPTUAL SCALING OF ROOM REVERBERATION.

Pavel Zahorik

Department of Psychological and Brain Sciences, University of Louisville, Louisville, KY 40292

UNIVERSITY of LOUISVILLE

dare to be great

ABSTRACT

Recent evidence suggests that reverberant energy can provide listeners with important spatial information regarding the distance of a sound source. However, relatively little is known about the perceptual attributes of the reverberation itself and how these attributes may be related to physical properties of the environment that also potentially impact perceived spatial location. Here perceived similarity among 15 reverberant rooms simulated using virtual auditory space techniques was examined. Room size and surface absorption properties were varied, along with aspects of the virtual simulation including the use of individualized head-related transfer function (HRTF) measurements and properties of the room acoustic simulation. Seven listeners rated perceived similarity on a 100-point scale between all possible pairs of simulated rooms using a speech source signal. Multidimensional scaling techniques were used to estimate scales of perceived room reverberation. Although the resulting scales were complex and somewhat unique to individual listeners, it is clear that the perceptual effects of manipulating properties of the reverberant sound are much larger than the effects due to either non-individualized HRTFs or non-optimal room simulation methods. [Work supported by NIDCD.]

MOTIVATION

Although the echoes and reverberation that result from the acoustic properties of rooms are known to affect the perceived spatial location of sounds in both direction (Hartmann, 1983) and distance (Zahorik, 2002), relatively little is known about the perceptual attributes of these room acoustic properties themselves. Here we seek to determine the perceptual attributes of small room acoustics using virtual auditory space techniques and multi-dimensional scaling.

METHODS

Listeners

Seven listeners (6 female) ages 18-29 years participated in the experiment. All had normal hearing, as verified by standard audiometric screening, and were experienced in sound localization tasks.

Binaural Room Impulse Response Measurements

Binaural room impulse responses (BRIRs) were measured for each participant in a single rectangular room (See Fig. 1) using methods described in (Zahorik, 2002). Binaural microphones (Sennheiser KE4-211-2) were used in a blocked-meatus configuration (Møller et al., 1995) and the sound source (Cambridge SoundWorks Center/Surround IV) was placed directly in front of the seated participant at a distance of 1 m. The room had dimensions of 5.7 × 4.3 × 2.6 m, and a broadband (125 - 4000 Hz) T_{60} of approximately 0.4 s.

Binaural Room Impulse Response Models

Simple models of binaural room impulse responses (BRIRs) were constructed using an image-model (Allen & Berkley, 1979) to simulate early reflections within a hypothetical rectangular room and a statistical model of the late reverberant energy. The direct-path and 500 early reflections were all spatially rendered using individualized head-related transfer functions (HRTFs) measured from 613 spatial locations surrounding the listener in an anechoic chamber. Each reflection was attenuated based on path-length, an average surface absorption coefficient (α), and the reflection order. Diffuse late reverberation was simulated using independent Gaussian noise samples for each ear shaped by decay functions derived from the Sabine equation in each of 6 octave-bands ranging from 125 - 4000 Hz. All room model processing was implemented using MATLAB® software. Fig. 2 displays the early portions of a measured BRIR from the left ear of a single participant, and a modeled BRIR designed to match the measurement room and conditions as closely as possible. A relatively good match may be observed.

METHODS (continued)

Fig. 1. The measurement room environment (5.7 × 4.3 × 2.6 m). The loudspeaker sound source, and seated listener location are also shown.

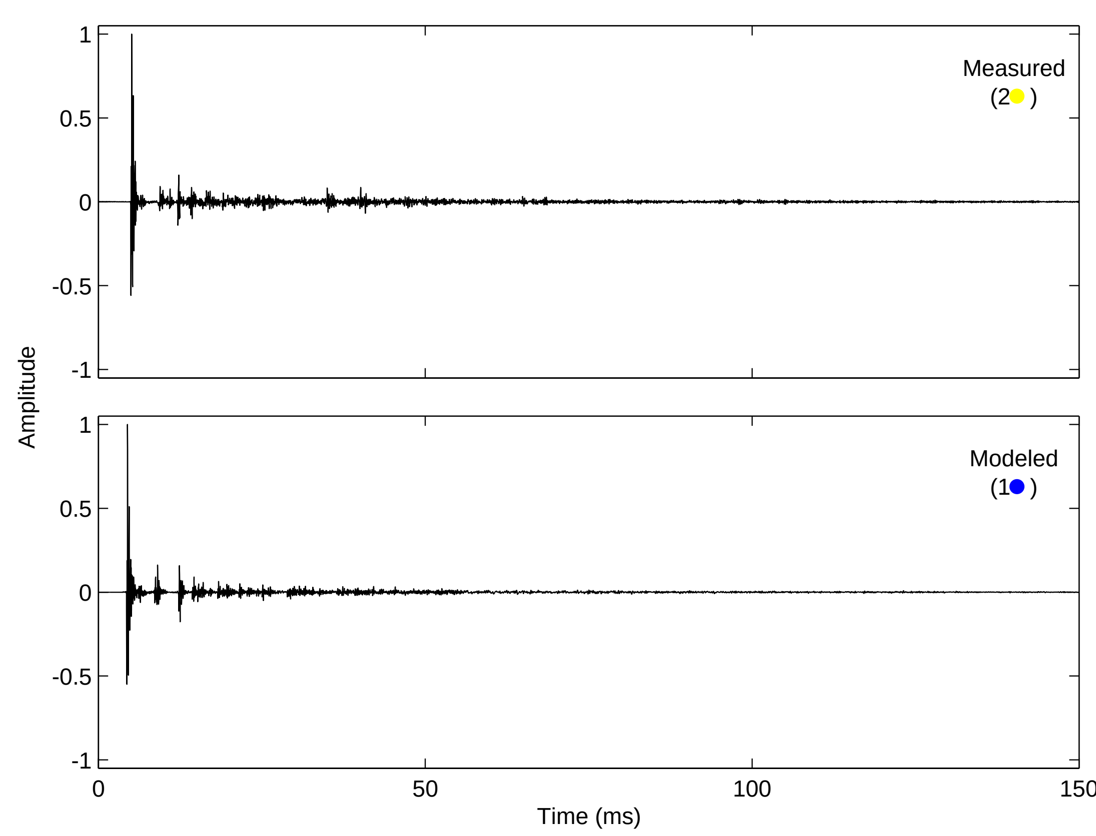

Fig. 2. Measured and modeled binaural room impulse responses (BRIRs) for listener SZI. Only the first 150 ms of the responses from the left ear are displayed.

Stimuli and Presentation Apparatus

Fifteen different stimuli were constructed from either measured or modeled BRIRs, as shown in Table 1. These stimuli included 3 different types of BRIR manipulations related to: room size, amount of late reverberant energy (controlled by adjusting the amount of surface absorption, α), and individualized vs. non-individualized HRTF rendering.

A high-quality speech sample (3.4 s duration) from a male talker recorded in anechoic space was used as the source signal for all stimuli. This signal was then convolved with either measured or modeled BRIRs (using MATLAB® software) and presented over Beyerdynamic DT-990-Pro headphones using Tucker-Davis Technologies equipment for D/A conversion (16-bit, 48 kHz) and analog gain control.

Design and Procedure

Participants listened to all possible pairs of different stimuli (210 total), presented with an ISI of approximately 1 s. Participants were told to rate the perceived (dis)similarity between each stimulus in the pair using a 100-point rating scale, ranging from 0 = "exact same" to 99 = "completely different." Participants were allowed to listen to the stimulus-pair as many times as they wished prior to making their rating response. No feedback was given to participants as to the type of trial or the nature of their responses. The experiment was run in blocks of 210 trials consisting of 1 set of all possible pairs of different stimuli. Participants required roughly 45 minutes to complete one trial block. Each listener completed 9 blocks of trials, resulting in a total of 1890 trials, or 9 similarity ratings for each stimulus pair.

RESULTS

1 ●	Best modeled BRIR (own HRTFs)
2 ●	Measured BRIR (own HRTFs)
3 ●	Modeled BRIR (other HRTFs 1)
4 ●	Measured BRIR (other HRTFs 1)
5 ●	Modeled BRIR (other HRTFs 2)
6 ●	Measured BRIR (other HRTFs 2)
7 ●	Modeled BRIR (own HRTFs, room size 0.6x)
8 ●	Modeled BRIR (own HRTFs, room size 1.5x)
9 ●	Modeled BRIR (own HRTFs, room size 2x)
10 ●	Modeled BRIR (own HRTFs, room size 5x)
11 ●	Modeled BRIR (own HRTFs, late reverb α = 0.05)
12 ●	Modeled BRIR (own HRTFs, late reverb α = 0.1)
13 ●	Modeled BRIR (own HRTFs, late reverb α = 0.3)
14 ●	Modeled BRIR (own HRTFs, late reverb α = 1)
15 ●	Modeled BRIR (own HRTFs, ANECHOIC)

Table 1. Stimulus conditions with measured and modeled binaural room impulse responses (BRIRs). Broadband (125 – 4000 Hz) T_{60} and interaural cross-correlation (*IACC*) value are also shown.

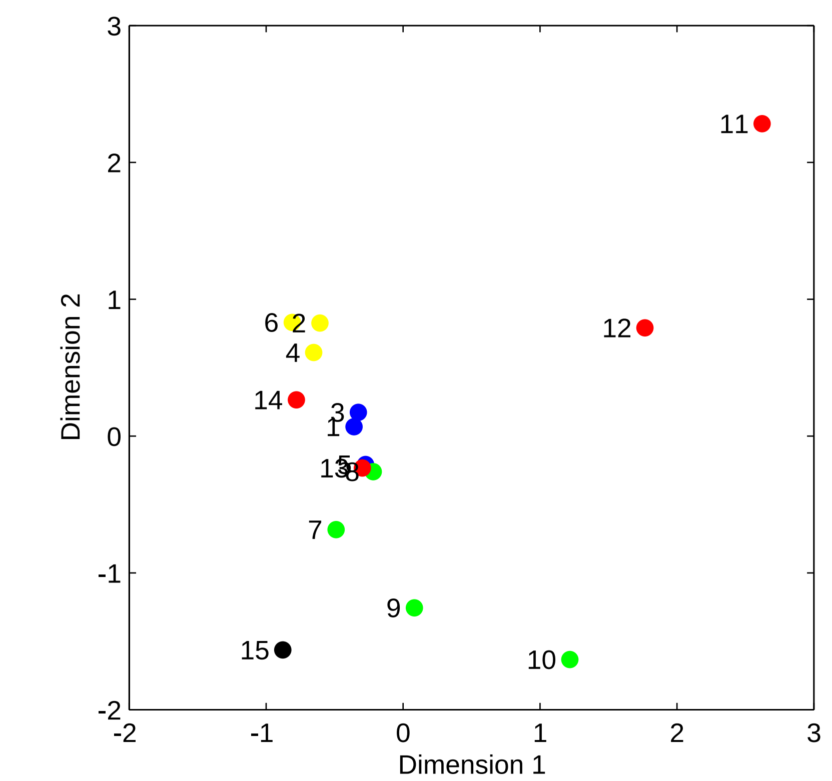

Fig. 5. Derived 2-dimensional scale of perceived room similarity (see Table 1 for stimulus number coding).

A multi-dimensional scaling analysis was performed on the average perceived similarity ratings from each listener. This analysis (INDSCAL, implemented using SPSS® software) allowed for a scale of perceived room similarity to be determined, as well as a characterization of each individual listener's use of the resulting scale. Here the scale of perceived similarity in room acoustics was determined to be 2-dimensional, since solutions with higher dimensionality did not account for a significantly greater proportion of the total variance (R^2), as shown in Fig. 3. The 2-dimensional solution also resulted in high R^2 values associated with each individual listeners' data, as shown in Fig. 4. Fig. 5 displays the 2-dimensional scale of perceived room acoustics. Dimension 1 was found to be highly correlated (r = .92) with T_{60}, and therefore likely represents a perceptual quantity related to reverberation time. It is less clear what perceptual quantity is expressed by Dimension 2 of the solution. Because Dimension 2 is moderately correlated with *IACC* (r = -.55), it may represent a spatial aspect of the perceived sound, perhaps related to image size or diffuseness.

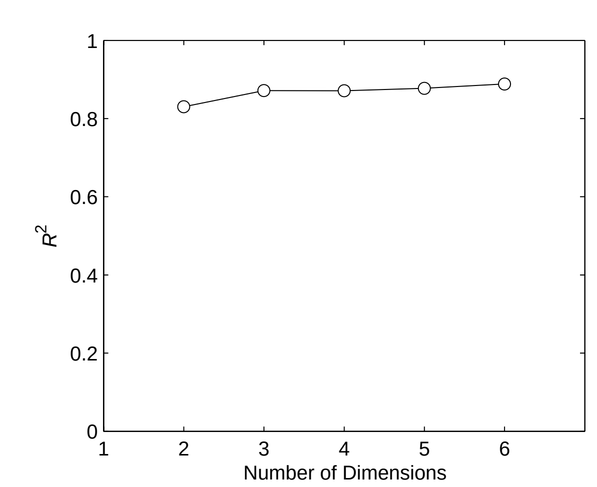

Fig. 3. Proportion of variance accounted for (R^2) as a function of scaling solutions dimension.

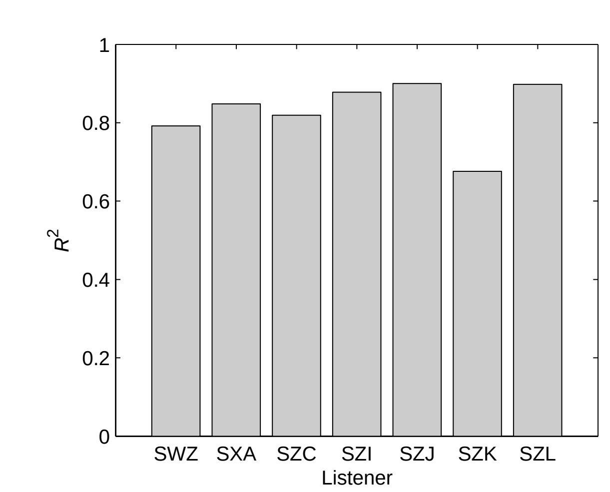

Fig. 4. Proportion of variance accounted for (R^2) in the 2-dimensional scaling solution as a function of individual listener.

It is also clear from this scaling solution that the modeled BRIRs do not exactly match the percepts elicited by the measured BRIRs. In general, the modeled BRIRs appear lie slightly higher on Dimension 1 and slightly lower on Dimension 2, which suggests that the modeled BRIRs are perceived as slightly more reverberant and slightly less diffuse than the measured BRIRs. It may also be concluded that measured and modeled BRIRs to not differ along any other perceptual dimensions, given the 2-dimensional nature of the solution demonstrated here. Of further note is the close proximity of modeled BRIRs using individualized HRTFs to those using non-individualized HRTFs. This suggests that the perceptual effects of manipulating properties of the reverberant sound are much larger than the effects due to the potentially degraded spatial rendering of the direct-path and early reflections with non-individualized HRTFs. The relative importance of the dimensions in the solution for each listener are shown in Fig. 6. The majority of listeners tended to place greater importance, or weight, on Dimension 1, and relatively less weight on Dimension 2. This suggests that for most listeners, aspects of room reverberation time are the primary means for judging the similarities between rooms. Some listeners however, (e.g. SZK) appear to base their judgments of room similarity primarily on the perhaps spatially oriented aspects of Dimension 2.

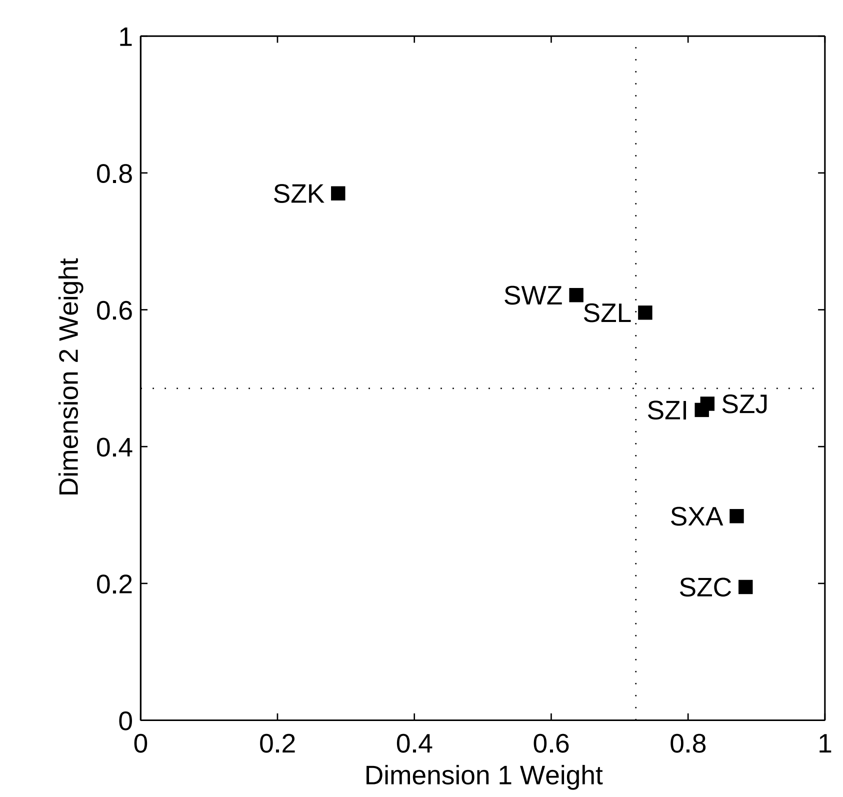

Fig. 6. Individual listener weights for the 2-dimensional scaling solution. Dotted lines indicate the average weights for each dimension.

CONCLUSIONS

- Listeners base judgments of room similarity on only two perceptual dimensions.
- Dimension 1 is highly correlated with reverberation time (r = .92).
- Dimension 2 is moderately correlated with IACC (r = -.55).
- Most listeners base their judgments of room similarity primarily on reverberation time (Dimension 1), although relatively large individual differences were observed.
- Effects of spatial rendering quality (individualized HRTFs) are small, which has important implications for virtual auditory display and room auralization applications (Kleiner et al., 1993).

Thanks to Jen Junion-Dienger for her assistance in data collection.
Work supported by NIH-NIDCD (R03 DC005709).

REFERENCES

Allen, J. B., & Berkley, D. A. (**1979**). Image method for efficiently simulating small-room acoustics. J. Acoust. Soc. Am., **65**, 943-950.

Hartmann, W. M. (**1983**). Localization of sound in rooms. J. Acoust. Soc. Am., **74**, 1380-1391.

Kleiner, M., Dalenbäck, B., & Svensson, P. (**1993**). Auralization - An overview. J. Audio Eng. Soc., **41**, 861-875.

Møller, H., Sørensen, M. F., Hammershøi, D., & Jensen, C. B. (**1995**). Head-related transfer functions of human subjects. J. Audio Eng. Soc., **43**, 300-321.

Zahorik, P. (**2002**). Assessing auditory distance perception using virtual acoustics. J. Acoust. Soc. Am., **111**, 1832-1846.