

Sound Localization in Partially Updated Room Auralizations

Samuel W. Clapp and Bernhard U. Seeber

Audio Information Processing, Technische Universität München, 80333 München, samuel.clapp@tum.de

Introduction

Room auralization is an important and versatile tool with many applications in both psychoacoustic research and in the design of new spaces (or renovation of existing spaces) where the acoustics play an important role [1]. Creating a room auralization requires a room impulse response, which contains the information on how the room transforms a sound signal between a given sound source position and a given receiver position.

Room impulse responses can be obtained either through measurements in a real room or through computer modeling. With computer models it is simple to adjust aspects of a room that in reality would require extensive renovations, such as changing surface materials or the positions of walls and ceiling. Due to the increasing power of modern computers, a developing area in room auralization is the implementation of real-time capabilities, where source and receiver positions can be changed, and the room impulse response updated, while the simulation is running. This requires the efficient calculation of updated room impulse responses to produce a realistic impression of movement for the listener.

The room simulations in this study are based on the image source method, where the spatial and temporal information of individual sound reflections is determined by geometrically reflecting the source position about the wall surfaces, which was developed for arbitrary room geometries by Borish [2]. An N th-order reflection is one that is reflected over N surfaces before reaching the receiver. Computation of the image sources becomes more costly with increasing order N , as the number of possible image sources increases geometrically.

The main question asked in this study is how much of the room impulse response must be updated, in order for a new source position to be localized to the same location as if a complete new room impulse response had been calculated for that position. This question is relevant for real-time room acoustic simulation with moving sources, as it may be possible to only calculate a room impulse response up to a given order of image sources within the time dictated by the latency requirements.

Methods

Simulated Room

A rectangular room was used for all test conditions in this study. The dimensions of the room were 5 x 9 x 2.3 meters, with a volume of 103.5 cubic meters. Frequency-dependent absorption coefficients based on real materials were applied to the room surfaces: heavy carpet on concrete for the floor, and unglazed, painted brick for the walls and ceiling. The broadband reverberation time was approximately 1 second.

The design of this room was chosen so that the receiver would encounter a large number of strong early reflections.

Three different linear source trajectories, located at different distances from the receiver, were used to simulate the impulse responses. Sources were simulated with the directivity of a human speaker, after the measurements of Flanagan [3]. The speaker was oriented along each trajectory line. A floor plan view of the receiver position and source trajectories is shown in Figure 1.

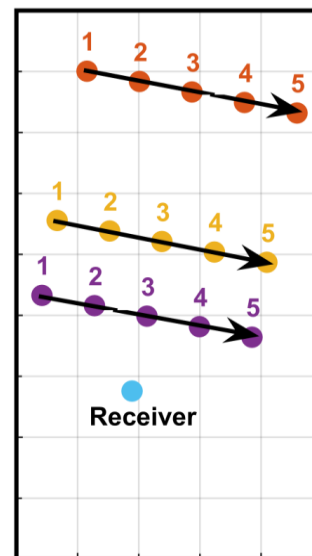


Figure 1. Floor plan view of simulated room showing source trajectories and receiver position.

The acoustic simulations were generated using the image source method by Borish [2]. Image sources were calculated up to order 200, to model the high reflection density of late reverberation. A 10% temporal jitter was applied to image sources beginning at 5th order, to simulate the effects of diffusion.

Test Environment and Rendering

This study was conducted in the Simulated Open Field Environment (v3) at the Technische Universität München [4]. The environment consists of 96 small loudspeakers arranged in a ring of 1.2-meter radius, with an angular spacing of 3.75 degrees between loudspeakers. Auralizations were rendered using the nearest-loudspeaker method, with each reflection mapped to the loudspeaker closest to its angular position. Reflections from outside the azimuthal plane were mapped to the loudspeaker located nearest the cone of confusion on which the elevated reflection sits. Localization judgments were solicited with a laser pointer controlled with a track ball (*ProDePo*-method) [5]. The laser pointer projects onto a white paper ring that sits just above the loudspeakers.

Test Conditions

All 15 source positions were simulated up to the maximum image source order (200). Six different source movement scenarios were considered along each trajectory, as shown in Figure 2. For each source movement scenario, a “hybrid” auralization was created using image source locations up to a certain order from the “new” position, and all higher order image sources from the “old” position. The “cutoff” image source order was one of four values: 0 (i.e. direct sound from the new position, all other reflections from the old position), 1, 3, or 7.

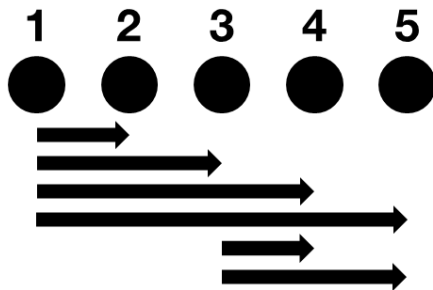


Figure 2. Six source movement scenarios considered along each trajectory.

Test Format

For each trial in this test, a stimulus was played, after which the laser pointer was activated, and test subjects could move it to the direction from which they heard the sound source originate. The test was performed in complete darkness, to avoid the possibility of subjects using visual anchors to make their localization judgments. Subjects were free to move their heads at all points during the test. When the laser was activated, it was positioned within a region of ± 20 degrees around the “true” geometric source position.

Before beginning the main portion of the test, each subject completed a training session of 30 trials, with no feedback given, to become familiar with the testing environment and methods. Each subject completed 8 runs of the experiment. The total time for completing all runs was around 90-100 minutes.

Results and Discussion

Selected results for the trajectory located furthest from the receiver are shown in Figure 3. The plots show the localization results (medians and upper and lower quartiles) from two source movement scenarios from this trajectory. Both start at Position 1 and move to Position 2 (right side of the plot) or 3 (left side). The dotted lines show the azimuthal angle corresponding to the old position (i.e. Position 1) in relation to the new position. The responses are plotted in degrees relative to the median of the “Total” condition (i.e. the most accurately simulated, up to order 200).

In both cases, for the 0th order update, significant localization errors were made in the direction of the old source position. For the Position 1 to 2 scenario, the median response was 7.5 degrees to the left as compared with the “Total” condition, a 100% localization error that indicates subjects were localizing directly to the old source position.

For the Position 1 to 3 scenario, the old source position was located approximately 19 degrees to the left of the new source position, while the median response was approximately 14 degrees to the left, not a 100% error, but still substantial.

Because this room has a relatively low ceiling of 2.3 meters, many early reflections reach the listener due to ceiling and floor bounces. These reflections originate from the same azimuthal location as the source. Thus, for an update of order 0, there is still a lot of sound energy originating from the direction of the “old” source position, which clearly has an effect on listeners’ localization judgments.

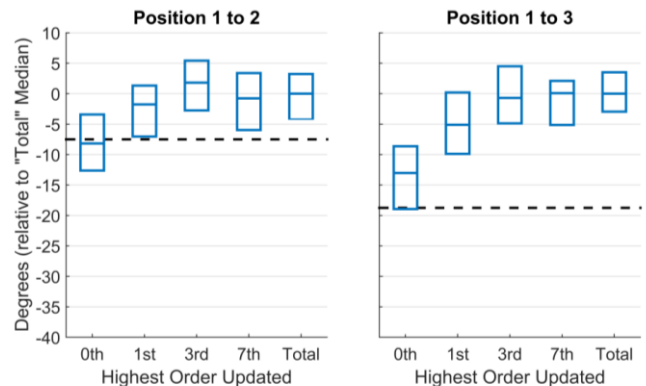


Figure 3. Localization results for far trajectory, with source movement scenarios starting at Position 1 and moving to Position 2 (left) and 3 (right).

Both source movement scenarios show the same overall trends for the higher update orders. 1st order update conditions also show an error in the direction of the original source position, although it is not as large. 3rd and 7th order update conditions exhibit median localization judgements that are within 1-2 degrees of the “Total” condition. This seems to indicate that for this particular room, a 3rd order update is sufficient to achieve similar median azimuthal localization results as would be obtained with a complete update to order 200. (However, some differences exist in the variance and distribution of responses.)

The nearer source trajectories generally showed the same trends, but with lower errors for the low update orders. However, there were two notable exceptions. The first occurred with the Position 3 to 5 scenario, the results for which are shown in Figure 4. Here there is around 100% localization error for both trajectories, which is not seen in the other scenarios. This is most likely due to the particular geometry of this condition. Because of the low ceiling and the fact that Position 3 is much closer to the receiver than Position 5 (for these nearer trajectories), the ceiling and floor reflections from the “old” position (Position 3) actually reach the listener before the direct sound from the “new” position (Position 5). Therefore, in these conditions, listeners are actually localizing “correctly” as predicted by the precedence effect. This represents a scenario that must be very carefully considered in a real-time auralization system.

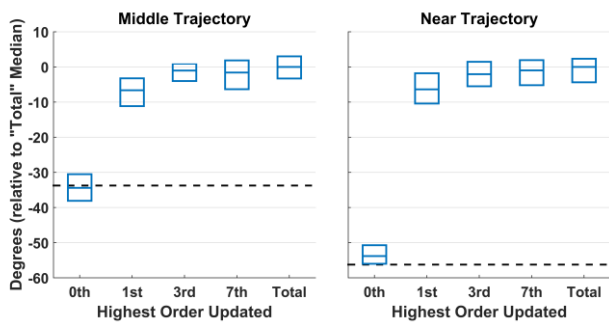


Figure 4. Localization results for Position 3 to 5 source movement scenario for middle trajectory (left) and nearest trajectory (right).

The second exception occurred with the Position 1 to 5 scenario, the results for which are shown in Figure 5. Here, the individual responses for the 0th order update are shown, rather than the medians and quartiles. From this it can be seen that listeners may have perceived a split auditory image, with a number of responses at the old source location, and a number of responses at the new source location (Subjects could only give a single response with the laser pointer and were not given special instructions on how to respond if they perceived multiple sources). These results may also be explained by the particular geometry of this condition. The old and new source positions are separated by large angles – 67.5 degrees for the middle trajectory and 112.5 degrees for the nearest trajectory. The direct sound will arrive first from the new position, followed by several strong early floor and ceiling reflections from the old position. However, due to the large angular separation, a split image is perceived, rather than a single, fused source.

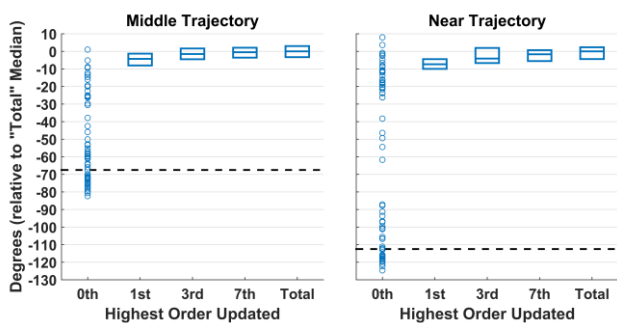


Figure 5. Localization results for Position 1 to 5 source movement scenario for middle trajectory (left) and nearest trajectory (right).

Acknowledgements

This work was supported by the Bernstein Center for Computational Neuroscience Munich, BMBF 01 GQ 1004B.

References

- [1] M. Vorländer, *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*, First ed. Berlin, Germany: Springer-Verlag, 2008.
- [2] J. Borish, "Extension of the image model to arbitrary polyhedra," *J. Acoust. Soc. Am.*, vol. 75, no. 6, pp. 1827–1836, 1984.

- [3] J. L. Flanagan, "Analog Measurements of Sound Radiation from the Mouth," *J. Acoust. Soc. Am.*, vol. 32, no. 12, pp. 1613–1620, 1960.
- [4] B. U. Seeber, S. Kerber, and E. R. Hafter, "A System to Simulate and Reproduce Audio-Visual Environments for Spatial Hearing Research," *Hear. Res.*, vol. 260, no. 1–2, pp. 1–10, 2010.
- [5] B. Seeber, "A New Method for Localization Studies," *Acta Acust. united with Acust.*, vol. 88, pp. 446–450, 2002.