Interactive Sound Propagation and Rendering for Large Multi-Source Scenes (Supplementary Material)

Carl Schissler and Dinesh Manocha
University of North Carolina at Chapel Hill

In this supplementary document, we present further results and analysis related to our algorithm and implementation:

Comparisons with the ground-truth measurements of room acoustic parameters *EDT*, C_{80}, D_{50}, G, and *TS* for the Elmia Round Robin benchmark.

Additional comparisons with the offline ODEON™offline acoustics system for the Elmia Round Robin benchmark.

A visualization of the simplified model used for acoustic simulation in the Tradeshow scene.

1. COMPARISON TO MEASUREMENTS

Here we present further comparison to measured data for the Elmia Round Robin benchmark. The results for the *EDT* parameter are shown in Figure 1. This parameter corresponds to the *early decay time* of the impulse response, i.e. the time for the sound to decay by 60dB when the decay rate is measured for the first 10dB of the reverberant decay. We observe acceptable agreement with the measurements for the middle and upper frequencies, but larger errors can be seen for the 125Hz and 250Hz frequency bands. The average error across all listeners and all frequencies for the *EDT* parameter is 7.6%. However, this is slightly greater than the just noticeable difference (JND) of 5%.

The C_{80} parameter describes the *clarity* of the sound at the listener's position and is computed as the ratio of the sound energy that arrives in the first 80ms to the sound energy that arrives after 80ms. The results for this parameter are presented in Figure 2. The average error across all listeners and frequencies is 1.5dB which is larger than the JND of 1dB. Once again, significant differences between our results and the measured data can be seen for the 125Hz frequency band, with an error of 2.3dB for the 125Hz band.

The D_{50} parameter is similar to C_{80} and describes the *definition* of the sound at the listener's position. The parameter, stated as a percentage, is computed as the ratio of sound energy in the first 50ms of the IR to the total sound energy in the entire IR. The results, shown in Figure 3, are similar to those for C_{80}. The average error over all listeners is 7.9%, which is more than the JND of 5%.

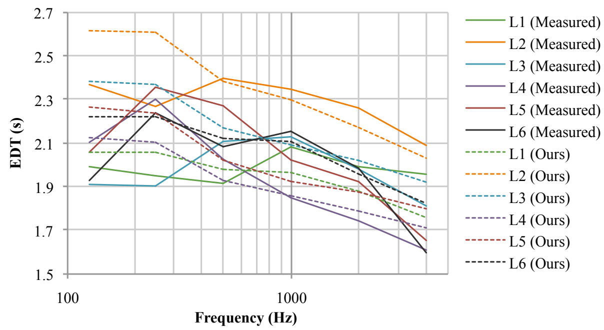

Fig. 1. This compares the ground-truth measurements in the Elmia Round Robin benchmark to our approach for the EDT (early decay time) room acoustic parameter.

The G parameter describes the strength of the sound at the listener's position. It is computed as a ratio of the total energy in the IR to the energy received by a listener 10m from the sound source in a free field. The results for this parameter are presented in Figure 4. There is good agreement at most listener positions and at higher frequencies, but there is also significant error at low frequencies for certain listeners (L_4, L_6). The average error across all listeners and frequencies was 1.1dB while the JND for the G parameter is 1dB.

The *TS* parameter, measured in milliseconds, is the center time of the impulse response and is computed as the time of the 1st moment of the sound energy in the IR. Figure 5 shows the results for this parameter. For some listener positions there is a close correspondence to the measured data (L_1, L_3, L_5), while the differences are greater for the other positions. Overall, the error for our method was 18.6ms, which is almost double the JND of 10ms.

DOI 10.1145/XXXXXXX.YYYYYYY
http://doi.acm.org/10.1145/XXXXXXX.YYYYYYY

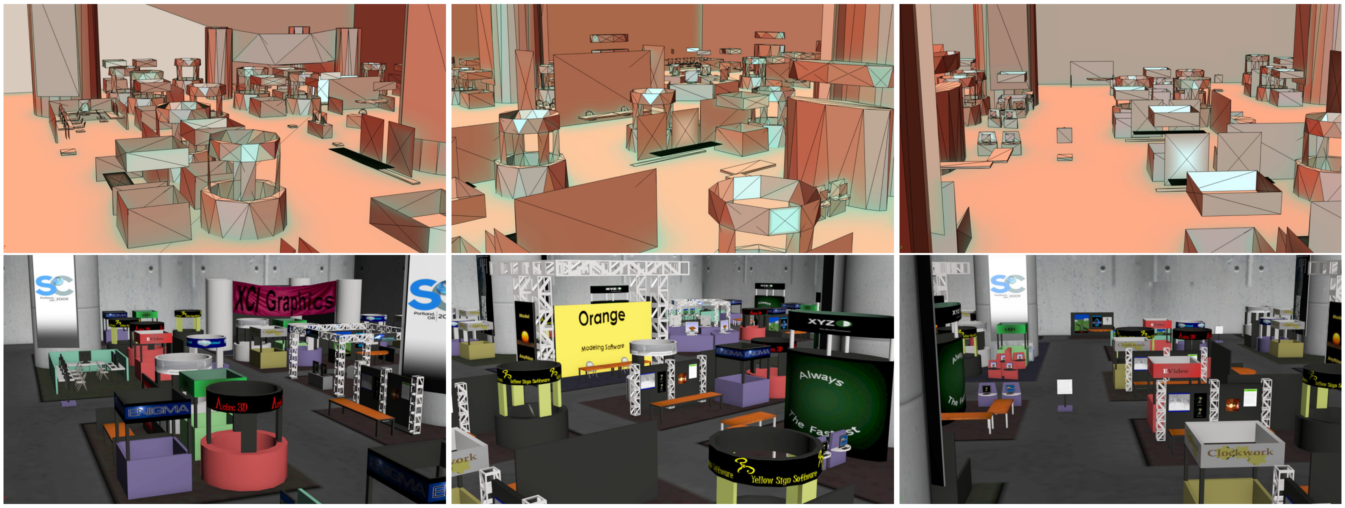

Fig. 6. A comparison between the simplified model used for acoustic simulation in the tradeshow scene and the model used for visual rendering.

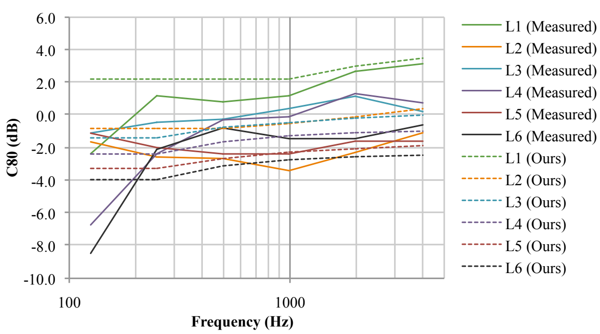

Fig. 2. This compares the ground-truth measurements in the Elmia Round Robin benchmark to our approach for the C_{80} (clarity) room acoustic parameter.

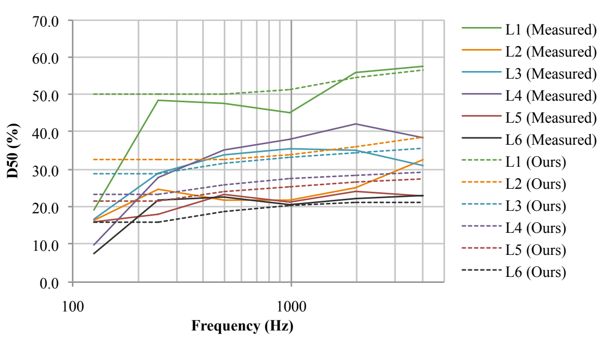

Fig. 3. This compares the ground-truth measurements in the Elmia Round Robin benchmark to our approach for the D_{50} (definition) room acoustic parameter.

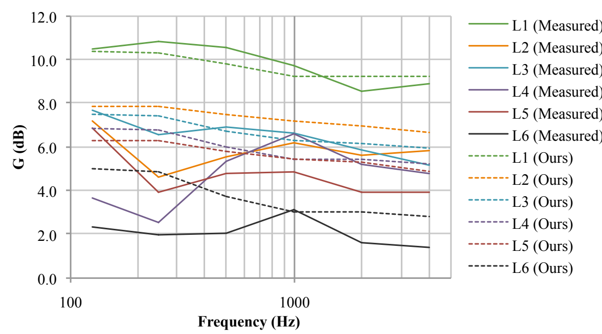

Fig. 4. This compares the ground-truth measurements in the Elmia Round Robin benchmark to our approach for the G (sound strength) room acoustic parameter.

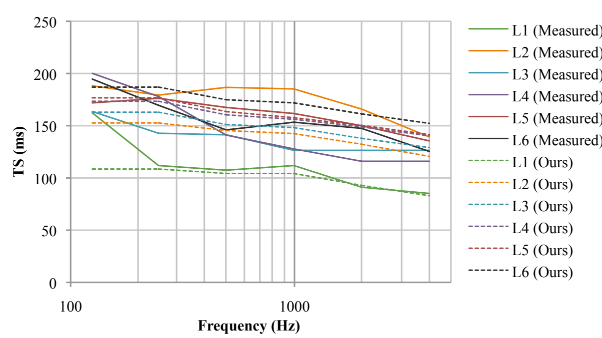

Fig. 5. This compares the ground-truth measurements in the Elmia Round Robin benchmark to our approach for the TS (center time) room acoustic parameter.

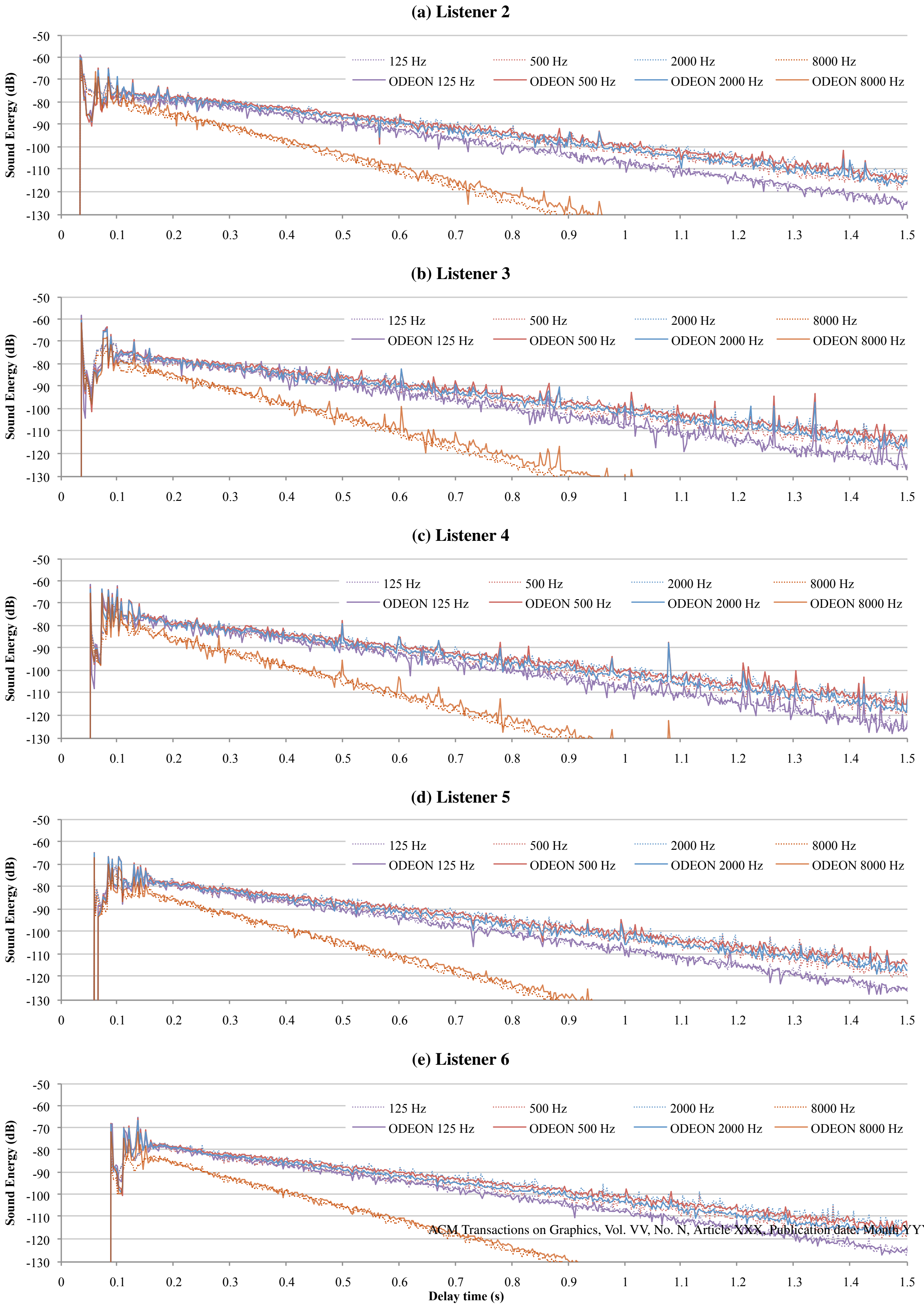

Fig. 7. This compares the sound energy decay for the 125Hz, 500Hz, 2kHz and 8kHz frequency bands of our system with ODEON™ for listener positions L_2 – L_6 in the Elmia Round Robin scene.

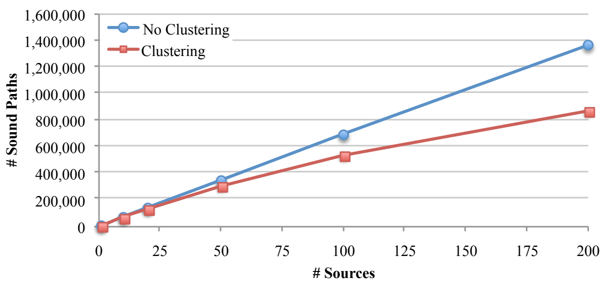

Fig. 8. The number of propagation paths (and the load on an audio rendering algorithm) for a simulation increases linearly with the number of sources. We show the total number of paths computed for the Tradeshow scene with up to 200 sound sources and with clustering enabled or disabled. Millions of paths are generated for large numbers of sources. We use a combination of clustering, propagation, and rendering algorithms to handle such complex scenarios.

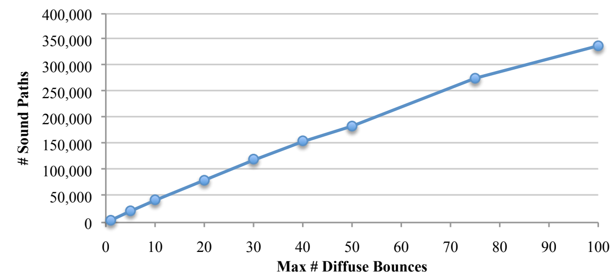

Fig. 9. We highlight how the number of propagation paths increases with the maximum diffuse reflection order on the Sibenik benchmark with 18 sources. Computing high-order reflections results in many more propagation paths, which are rendered in realtime by our hybrid audio rendering algorithm.