Supplementary Material

Precomputed Wave Simulation for Real-Time Sound Propagation of Dynamic Sources in Complex Scenes

Nikunj Raghuvanshi, John Snyder, Ravish Mehra, Ming Lin, Naga Govindaraju

Error Analysis in Living Room Figures 1, 2, and 3 analyze error at three different listener locations in the living room scene. See also Figure 7 in the main paper, which plots error at a fourth listener location. Error is shown in two parts: first (top of each figure), from encoding alone using our representation of peak times/amplitudes and a frequency trend, and second (bottom of each figure), from both encoding and interpolation at a listener location midway between the simulated grid points. Errors are computed with respect to a higher-frequency (4kHz) reference simulation, and so include frequencies beyond our "working" simulation which is bandlimited to 1kHz. Overall, compression error is low, often near the threshold of audibility, while total error (interpolation + compression) is somewhat higher but still reasonably small. In all cases, our method of interpolating peak times and amplitudes better preserves the high-frequency content of the impulse response than does straightforward linear interpolation of the signals. Not only is our result more accurate, but it also avoids artifacts from linear interpolation in which high sound frequencies are alternately preserved at the grid points and then attenuated between them as the listener moves.

"Shoebox" Experimental Results As a reference simulator, we have implemented the Image Source method for a simple rectangular shaped room. The walls are assumed to be purely specular and have no frequency-dependent absorption. However, atmospheric absorption is taken into account. In the paper, we have mentioned that above 5 kHz, frequencies are attenuated quite aggressively by the air. The spectrogram on the right of Figure 4 demonstrates this clearly.

Figure 5 shows further comparisons of the errors generated by our IR parametrization and interpolation approach. We have compared the Frequency Response obtained with the Image Source method for a source and listener location in the room with the response obtained with our approximations, namely, using a simulation limited to 1 kHz, parameterizing it with peak delays and amplitudes plus frequency trend, and interpolating it spatially. For this simple scene, Image Source method provides perfect interpolation, yielding the response at any desired point thus serving as a good reference. [1] Our results show very good agreement with the reference – while the maximum error is around 5 dB, as in all our previous tests, the average error hovers around 2-3 dB for the whole spectrum, up to 16 kHz. In contrast, linear interpolation performs far worse, underestimating the energy in all octaves above 2 kHz by about 7 dB.

[1] However, the corners of a room do produce diffracted scattering which is captured by our simulator but not the Image Source method. However, this can be safely disregarded by choosing high wall reflectivity and keeping the source and listener far from the corners, as we have done.

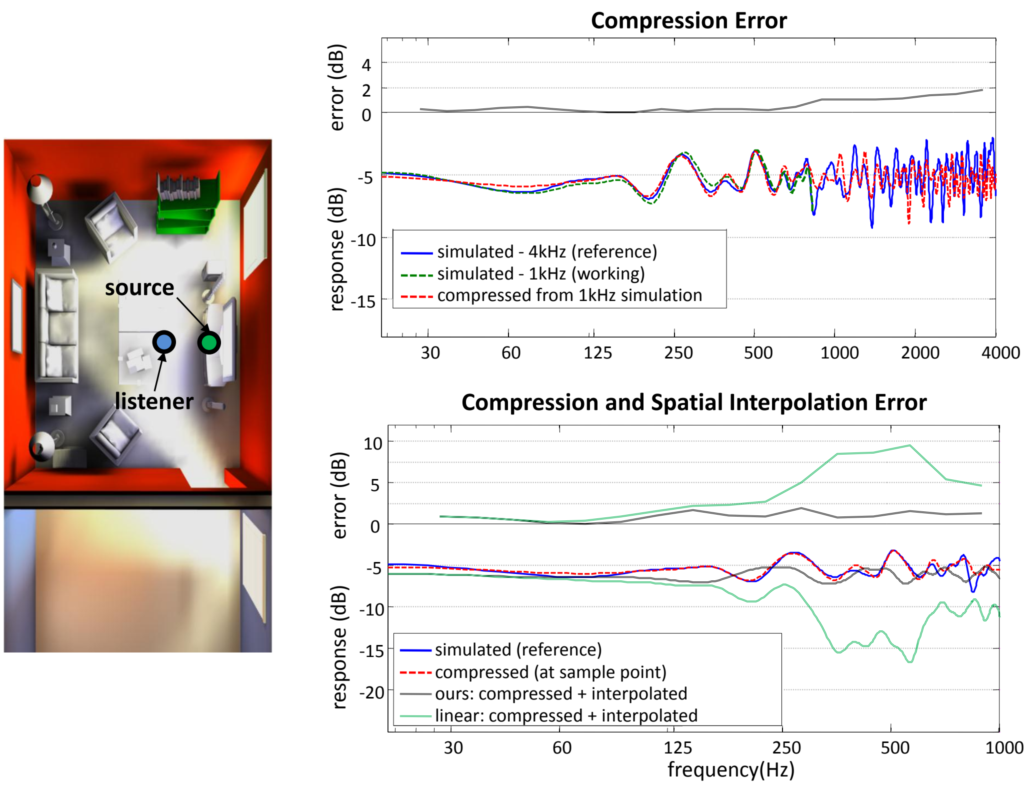

Figure 1: Error analysis: listener close to source. Our method matches the reference solution very closely, while linear interpolation yields substantial errors.

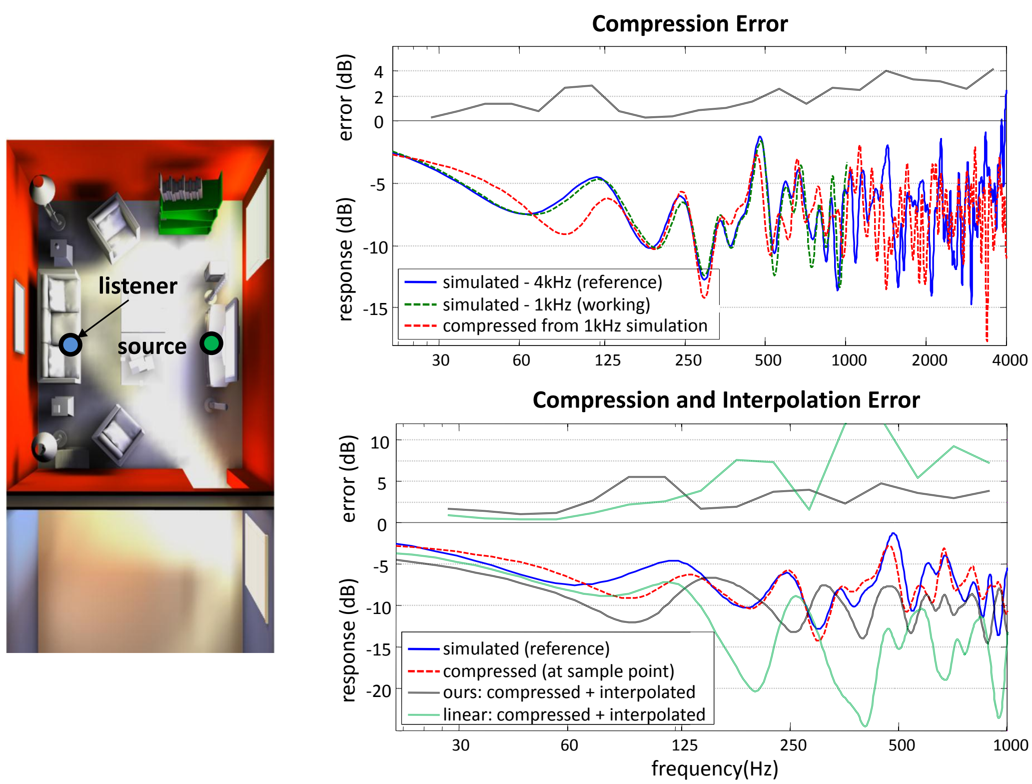

Figure 2: Error analysis: listener on couch. Compression error stays around 2dB till 1kHz and then increases to 4dB to 4kHz. Linear interpolation produces much more error.

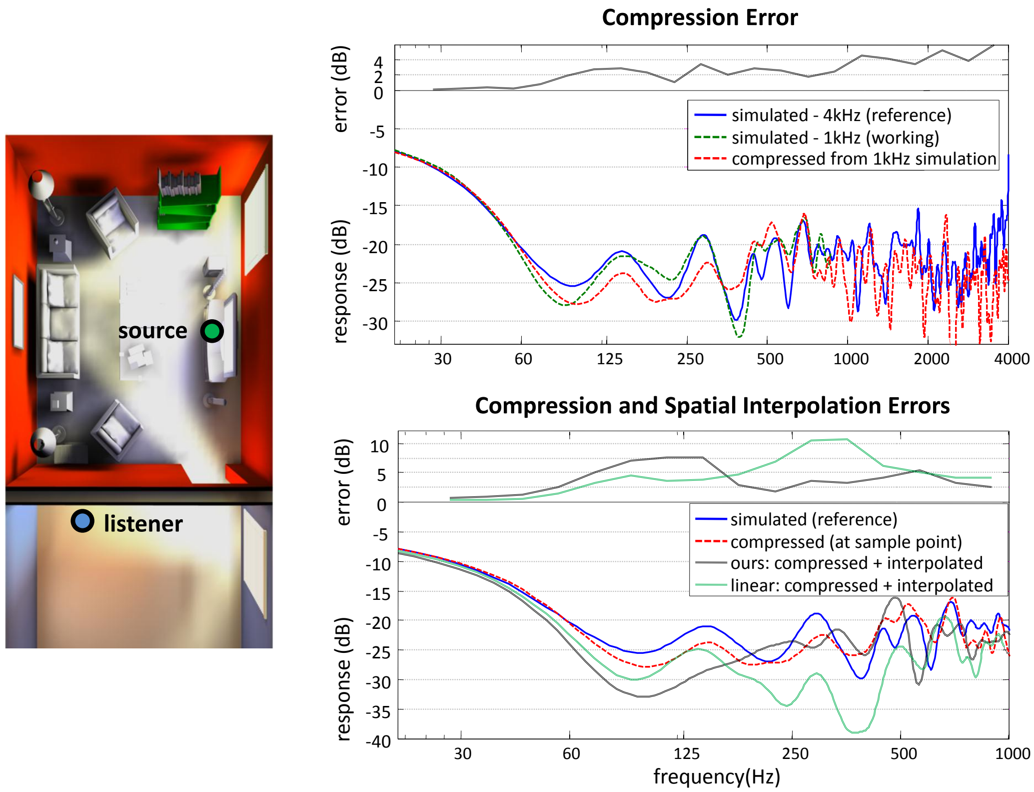

Figure 3: Error analysis: listener outside door. This is a highly occluded situation and a challenging case for current systems – Sound from the source undergoes multiple scattering in the room, then diffracts around the door to arrive at the listener. Such high-order effects are very hard to model convincingly. Notice the clear low-pass filtering in the frequency response which is also audible in the demo. Compression error lies between 2 to 4 dB, which is quite low. Spatial interpolation errors are higher, crossing 5 dB, but our technique still yields less error over all frequencies than linear interpolation.

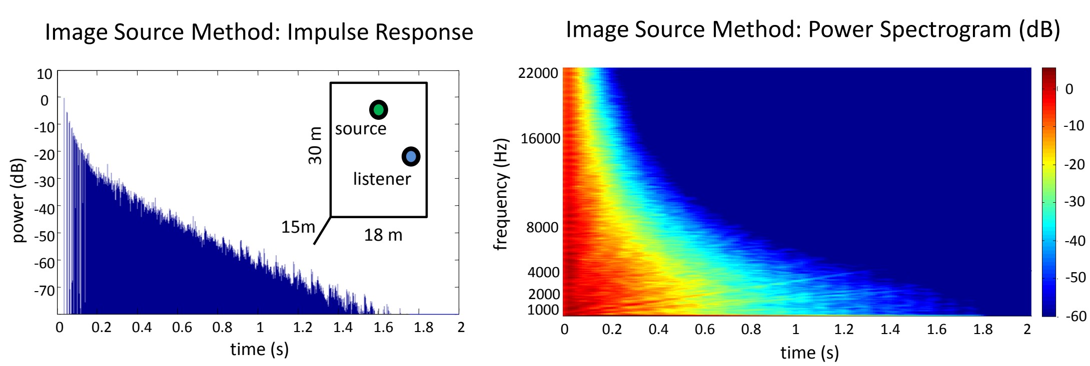

Figure 4: Effect of atmospheric attenuation The Image Source method was implementation for a simple "shoebox" geometry shown in the inset on the upper-right. Impulse responses were calculated assuming a pressure absorption coefficient of 0.15, frequency-independent absorption, but with atmospheric attenuation using the formulae given in ISO 9613-1 assuming an air temperature of 20 degrees centigrade and relative humidity of 40%. The spectrogram on the right clearly shows the aggressive attenuation of frequencies above 5kHz. In particular, a 10 kHz component sound would decay by nearly 20 dB after just 200 ms of propagation. In a real scene where material absorptions are also accounted for, this attenuation would be even higher.

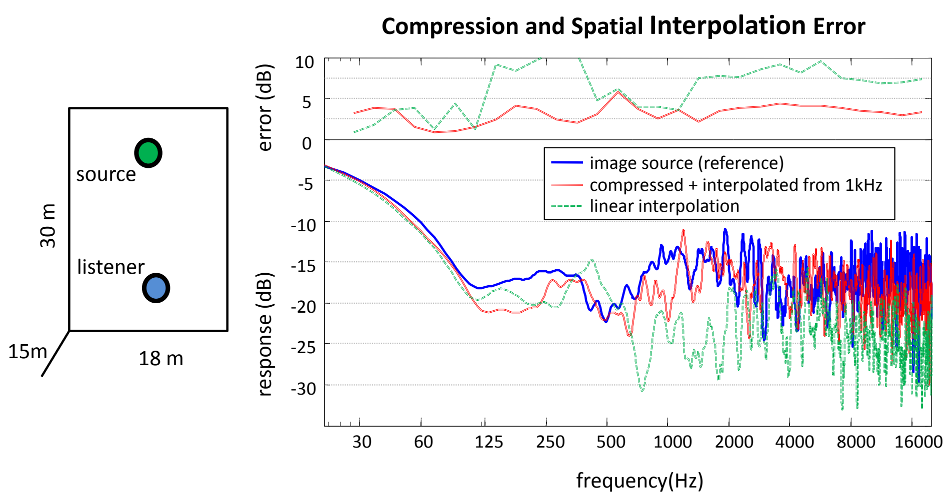

Figure 5: Comparison with Image Source Method The left image shows the locations of source and listener. A reference 200 ms long IR was generated with the image source method. The corresponding frequency response is plotted against the corresponding result obtained with our technique by using a bandlimited simulation to 1kHz, parameterizing the IR and then performing spatial interpolation on it. Our results show good agreement with the reference solution. From the error plot on the top, we can see that the maximum error is about 5 dB and the average error over the full frequency range till 16 kHz is 2-3 dB. In contrast, linear interpolation gives much larger errors.