Room Impulse Response Synthesis with Device Diffraction via Image Source Method and Finite Element Analysis

Aidan Meacham[1, a)] **and Andrew Unruh**[1]

Knowles Intelligent Audio, Mountain View, CA 94043, United States

aidan.meacham@knowles.com,

andy.unruh@knowles.com

(Dated: 27 October 2017)

Abstract:

A method to approximate the soundfield on the surface of a device was developed by combining the image source method and anechoic finite element simulations. Impulse responses from finite element simulations of plane waves incident on the device can be assembled like image sources to derive transfer functions from sources in space to receivers on the surface. The direction of each image source is matched with a simulated plane wave, the corresponding impulse response is delayed and attenuated according to the image source's time of travel and reflection path, and all are summed to create a synthesized reverberant impulse response.

[a)] Author to whom correspondence should be addressed.

1. Introduction

Geometrical and wave-based approaches to reverberant room acoustics are well-understood and have applications in a variety of domains, but also have structural downsides. While geometric algorithms (such as ray tracing or the image source method) tend to work well for macroscopic acoustic situations, they are typically unable to account for edge diffraction or other small-scale phenomena. Those features become increasingly relevant when approximating the soundfield on the surface of, say, a mobile phone or IoT device. On the other hand, wave propagation strategies (such as finite element analysis or finite-difference time-domain methods) are particularly well suited for modeling minutiae, but for large spaces at any but the lowest frequencies, the technique becomes computationally untenable.

Both of these techniques are well known, so we will not describe more than the pertinent details for the hybrid method. Finite element analysis is an approach where the problem to be solved is quantized into small volumetric regions (the "finite elements") that are assigned acoustic properties that correspond to their real-life counterparts. Pressure waves that propagate through these elements can diffract around edges, lose energy to absorptive materials, and reflect off of hard surfaces. The size of the elements is chosen relative to the wavelength of the highest frequency that must be simulated. One relevant concept is that of a perfectly matched layer (PML) which, when applied to a boundary, eliminates reflections of incident waves.[1] Using this technique, a free-floating device in an anechoic chamber can be modeled by a partially absorptive surface (representing the device) surrounded by a layer of air, and with a PML applied to the outer surface of the air boundary. Thus, a plane wave entering the problem domain will propagate through the air, impact the device, diffract around it, and continue propagating to the PML where it is absorbed, as if it were continuing into space with no change in impedance. The benefit of this technique is that it allows us to put a boundary on the size of the problem to be solved that isn't reflective (such as a wall), which can be much smaller than the size of a room or anechoic chamber, for example. Since reducing the number of elements also reduces the computation time, making the problem small and anechoic greatly speeds up the solving process. Many solvers operate in the frequency domain, so deriving an impulse response from a given source to a point on a device is a matter of simulating a range of frequencies, assembling the spectra, and performing an inverse Fourier transform. The resulting impulse responses should be relatively short given the anechoic nature of the simulations, but it is important to allow enough samples to reach a desired level of decay. In this case, the number of frequencies that are solved (up to the highest frequency of interest) correspond to half of the number of samples in the impulse response.

The image source method models reverberant impulse responses in a given space by summing the effects of individual virtual sources whose positions are determined by reflecting the initial source across each boundary in the room.[2] For a rectangular space, higher order reflections are calculated by tessellating the room and reflecting the previous order of images across the newly-formed boundaries. This process is repeated until a desired level of decay (according to the distance of image sources and number of reflections) has been reached. To assemble the reverberant impulse response, each virtual source contributes a copy of the direct path's impulse response, which is delayed and attenuated according to the path length it must travel, and is additionally filtered by each surface with which it interacts.

In this paper, we present a method that takes advantage of aspects of both of these approaches and combines them to derive reverberant impulse responses that accurately model the soundfield on the surface of a device of interest. We will discuss the methodology, tradeoffs between accuracy and speed, and evaluate its performance relative to measured data.

2. The Hybrid Method

Consider a device in the free field surrounded by a theoretical spherical shell centered on the device's centroid. Any sound energy that is incident upon the device from outside the shell must, at some point, pass through it. If we have a single point source in an anechoic environment with the device, the point at which energy first passes through the sphere lies along the line from the source to the center of the shell. Under the assumption that the point source is sufficiently far away, we may approximate its wavefront with a plane wave approaching from the direction of the source. Thus, by scaling and delaying a plane wave that is emitted from the point of intersection, we can approximate the soundfield produced by the point source on the device.

Now, let us consider discretizing all possible directions of incidence. This is the same thing as picking a set of n points on the spherical shell from which plane waves may emanate. For any source in space, we can approximate the resulting soundfield on the device by picking the discretized point on the shell that has the smallest angular difference with the direction of the source and performing the aforementioned scaling and delay. For each of the n points on the shell, we can precompute a finite element simulation that gives us the transfer function for a plane wave beginning at that point to every point on the surface of the device. In doing so, we have assembled a “basis set” that can approximate the soundfield produced on the device by acoustic sources in any direction. From this basis set, we can assemble impulse responses from an arbitrary number of sources in space while still observing the effect of diffraction around the device as a result of the finite element simulations. Thus, this approach can be utilized to match the large (and relatively evenly distributed) number of sources that make up an image source method calculation. It is convenient that the mechanics of the image source method justify the use of plane waves since the vast majority of the virtual sources will be very far away and a plane wave is a reasonable approximation of their wavefronts.

In a typical implementation of the image source method for an empty room, once the location, time of travel, and attenuation for each source has been calculated, unit impulse responses are scaled, delayed, and summed to create a synthesized reverberant impulse response. In our methodology, the concept is largely the same; however, the unit impulse responses are replaced with individual impulse responses from the precomputed dataset, where the IR to use for a given virtual source is chosen by the closest discretized direction of arrival before going through the same scaling, delaying, and summation.

One significant difference are the path lengths (which help to determine the delay and scaling for a given source) that are computed. In a typical image source method implementation, the only important path length is measured from source to receiver, but since the sphere of plane waves already have delay built into their propagation, we must also consider the path length from the source to the point of origin of the plane wave that has been selected. We then calculate the delay based on this intermediate distance and the speed of sound.

It is simple to adjust the amplitude of the plane wave's impulse response according to the distance of the point source. Since we cannot assume that the pressure at a particular point on the device is a function of distance to the source (as a result of diffraction), it doesn't make sense to attempt to match the amplitude of the plane wave and point source on the surface. Instead, we match their pressures as if they had reached the center of the sphere. We centered the sphere at the centroid of the device, so we assume that this selection will minimize error everywhere on the device by virtue of proximity. If a point source were defined to produce 1 Pascal at 1 meter, and was 1 meter away from the center of the sphere, then (trivially) the pressure at the center of the sphere would be 1 Pascal. Similarly, if a plane wave were defined to have constant amplitude 1 J/m^2, then the pressure at the center of the sphere would also be 1 Pascal. This suggests an attenuation factor (applied to the plane wave's amplitude) that achieves unity at 1 meter and matches the amplitude of the point source, which is known to be $1/r$. Thus, a reverberant impulse response assembled from as precomputed plane wave finite element simulations with the given amplitudes and attenuation factors is identical to measuring the transfer function from a point source in the room that produces 1 Pascal at 1 meter to a point on the device.

3. Discussion

3.1 Discretization of Direction of Arrival

The process of picking an optimal method for discretizing the sphere into regions that correspond to a particular direction of arrival is congruent with the problem of evenly distributing points on a sphere. The problem can be presented as minimizing the maximum angular error between a random point and its nearest neighbor, which corresponds well to the issue of picking the best point for a given virtual source in the context of an image source method model. While an analytic solution to the problem does not exist (for any number of vertices except the Platonic solids, $n = 4, 6, 8, 12, 20$), there are a number of practical methods for creating a "reasonably good" distribution. Two of the best candidates are Fibonacci lattices[3,4] (sometimes referred to as "golden section spiral spheres") and subdivided icosahedrons (or geodesic polyhedra), each of which have properties that make them preferable for certain situations and are entirely deterministic.

As a quick overview, "Fibonacci lattices" (as we will refer to them) can be constructed for any number of points n by evenly distributing them along the elevation and azimuth, where the difference in azimuthal angle between successive points is chosen to be highly irrational. One way of visualizing this method is drawing a spiral from one pole to the other and then evenly distributing points along the unwrapped line segment. Note that many implementations pick phi as the azimuthal constant, hence the name.

Alternately, one can construct a subdivided icosahedron by beginning with a regular icosahedron and iteratively subdividing faces and projecting them onto the unit sphere. The downside to this methodology is that the algorithm only produces a few usable values of n

given a particular number of iterations, but the advantage is that the resulting lattice has symmetry across all three axes. Note again that it is beneficial to pick a regular icosahedron with highly irrational side lengths, such that vertex coordinates are often chosen to include phi.

In either case, the primary metric for picking a particular number of points n is by picking an arbitrary average or maximum angular error between a random point on the sphere (which corresponds to an arbitrary point source in space) and its nearest neighbor. Psychoacoustically, the minimum audible angle (the angle at which the average listener can differentiate between sources in the azimuth of varying types and frequencies) provides a lower bound of one degree, but the accuracy necessary in a given context may be far lower.

A secondary consideration for the number of points is the ability to match the theoretical coherence of the diffuse field for two receivers inside the sphere. For the case $n = 1$, it is trivial to see that the single source will have a coherence of 1, which is a poor approximation of the diffuse field. As n increases, the coherence between the two receivers approaches the theoretical value, but again, an arbitrary metric must be adopted in order to bound the number of simulations necessary. In our experience, the number of sources necessary to satisfy this requirement is far lower than the number needed to guarantee a particular angular resolution.

3.2 Symmetries

If a device has some form of symmetry, then it may be possible to reduce the number of simulations necessary to provide a full basis set by mapping degenerate sources onto the minimum number necessary. Consider a device that can be approximated by a rectangular solid of uniform material that is centered at the origin and aligned with each axis. Because of its triplete axial symmetry, a plane wave incident on the device originating from a point in the first octant (+X, +Y, +Z) will have the same transfer function to a receiver in the second octant (-X, +Y, +Z) as the source-receiver pair where each X-coordinate has been reflected across the YZ-plane. This degeneracy can be exploited to reduce the number of sources that must be simulated to one octant, an eightfold reduction. Similar arguments can be devised for devices with rotational symmetry, while keeping the goal of a particular arbitrarily-chosen maximum angular error for a random source in space.

3.3 Plane Waves

It is important to note that the model as described is only effective for sources that are sufficiently far from the device such that a plane wave is a reasonable approximation of omnidirectionality. The simplest way to confirm that a source is sufficiently far away is to simulate a point source and a plane wave (that points at the device) originating at the same location and compare their response on the surface of the device. If the difference is less than a given tolerance everywhere on the device, then the plane wave is clearly sufficient. Otherwise, the impulse response for that image source can trivially be replaced with that of the simulated point source. The main factors at play here are (nominally) the size of the device, since a larger (or more complex) device will experience more diffraction, magnifying the difference in wavefronts, and the distance of the sources, as the spherical wavefront appears more and more planar the further away it becomes.

If one is expecting to have many sources that will exceed the approximation tolerance (as in sources close to the device), it may be beneficial to precompute many candidates in a similar fashion to the sphere of plane waves, but in our experience, these are relatively uncommon and precomputation can typically be eschewed in favor of "runtime" simulations for those few sources near enough to invalidate the plane wave approximation.

3.4 Complementary Extensions to the Image Source Method

Like any other image source method implementation, there are a variety of improvements that can be made with varying degrees of triviality in order to handle different room geometry or areas of absorption, non-point sources with a specified directivity pattern, and more complex systems for modeling frequency dependent (or even nonlinear) losses through the air and at boundaries. In any case, these extensions often simply modify the number or treatment of individual virtual sources, and can be directly applied to the methodology discussed here. Thus, any improvements that are imaginable for the image source method can be utilized with little modification in the context of a hybrid model.

4. Analysis

In this section, we will present a comparison of impulse responses derived synthetically using the hybrid method and impulse responses collected on a device in a controlled environment. The measurements were performed in an ETSI-qualified chamber with known frequency-dependent T_{60} and dimensions, and a simulation was performed with those known parameters as inputs. The location of the device, a rectangular solid with microphones embedded so as to be flush, was identical in the real and virtual environments, and was mounted in such a manner as to minimize additional reflections (which would not be present in the virtual model). We examined individual impulse responses from various source locations in the chamber to each microphone, and the corresponding point on the surface of the virtual device. For each IR and combinations thereof, we examined the band-limited IRs, band-limited energy-time curves, 1/3-octave smoothed frequency responses, relative levels (with a designated primary microphone matched at its peak), and inter-mic coherence (as compared to the theoretical coherence in a diffuse field for a given mic spacing). In all cases, the results were determined to be satisfactory.

5. Conclusion

We have presented a methodology for the creation of reverberant impulse responses combining the image source method and finite element analysis while preserving the best attributes of each approach to room acoustics. We have discussed a variety of decisions that must be considered in the course of implementing such an algorithm, and finally, have experimentally verified that the method provides satisfactory approximations of real impulse responses. Overall, this hybrid synthesis technique extends geometric and wave-based acoustics into new use cases that would not be possible with either approach alone.

References and links

[1]J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," Journal of computational physics **114**(2), 185–200 (1994).

[2]J. B. Allen and D. A. Berkley, “Image method for efficiently simulating small-room acoustics,” The Journal of the Acoustical Society of America **65**(4), 943–950 (1979).

[3]E. A. Rakhmanov, E. Saff, and Y. Zhou, “Minimal discrete energy on the sphere,” Math. Res. Lett **1**(6), 647–662 (1994).

[4]E. B. Saff and A. B. Kuijlaars, “Distributing many points on a sphere,” The mathematical intelligencer **19**(1), 5–11 (1997).