

Directional Receiver Modelling

Introduction

In the literature, two methods for modelling directional receivers in the digital waveguide mesh (DWM) are presented. This essay will compare and evaluate the two methods, and explain the choice of method for a tetrahedral waveguide auralisation model, and describe the implementation of this chosen method.

The Blumlein Difference Technique

The first method, described by Southern & Murphy (2007a) and Southern & Murphy (2007b), is designed with the goal of encoding the 3D sound-field by means of spherical harmonic decomposition to the second order, for use in ambisonic reproduction.

This method is based on sound intensity probe theory, which allows directional velocity components of a sound field to be captured using two closely-spaced omnidirectional receivers. In the DWM, each node has an associated pressure, so two closely-spaced nodes can be used to emulate a pair of omnidirectional receivers, and in turn used to generate directional velocity components. A full first-order B-format recording can be made using seven such ‘pressure sensors’, arranged in a cross shape, with a single node at the centre, and six nodes placed around it. A second-order recording is achieved by adding further pressure-sensing nodes in between the first-order nodes.

While this technique requires the modelling of multiple coincident directional microphone capsules (as would be used to produce a ‘real’ ambisonic recording), the only important orientations are the *relative* orientations of the capsules. The position and orientation of the microphone array is not important, as a post-processing decoding step will be used to extract the desired directional information. The technique put forward suggests aligning the pressure sensors with the axial directions of the mesh, resulting in directional components which are also aligned to the mesh. For this reason, the technique is not very flexible when directly generating directional components.

Additionally, this technique has some difficulty adapting to different mesh topologies. In a rectilinear mesh, the seven first-order nodes can be placed exactly, as their positions will coincide with nodes in the mesh. However, the second-order nodes will not fall exactly on existing nodes, and instead will need to be ‘snapped’ to the closest mesh nodes. This problem becomes worse in other mesh topologies, such as the tetrahedral mesh, where it may be that very few of the sensor nodes can be exactly placed in the desired positions.

It is suggested to increase the spatial sampling frequency of the mesh in order to facilitate more exact placement of sensor nodes (by reducing the grid spacing, the quantisation error in the sensor placement is also reduced). The example given by Southern & Murphy (2007b) uses a sensor spacing of 0.04 metres, giving a maximum valid sampling frequency of 4.25kHz. However, to allow accurate placement of these sensors, a grid with a spacing of 0.0027m is used. Such a grid would normally support a maximum valid sampling frequency of 176.4kHz for a single-sensor omnidirectional receiver. This mesh oversampling is very costly, as the computational cost of the 3D DWM increases cubically as the spacing of the mesh is reduced.

In conclusion, although promising for producing ambisonic recordings in rectilinear meshes where computational time is not a great concern, this method does not seem well-suited to producing fast, arbitrary directional recordings in the tetrahedral mesh.

The Intensity Calculation Technique

The second method is described by Hacıhabiboglu, Günel, & Cvetkovic (2010). This method is based around estimating the acoustic intensity at the output node of the DWM.

The technique is much more straight-forward than the technique discussed above. The pressure differences between the ‘output’ node and its surrounding nodes are calculated and placed into a column vector. This vector is multiplied by the inverse of a matrix containing the relative positions of the surrounding nodes to the output node and, the output of this calculation is a three-element vector containing an approximation for the derivative of pressure with respect to each of the three axial directions. The velocity for each time step is found by adding the pressure-derivative vector to the velocity at the previous time step, where the initial velocity is assumed to be zero. Intensity is found by multiplying this velocity by the pressure at the output node.

Under this scheme, the output of the waveguide mesh is actually two streams of information, describing the pressure and the intensity at the output. Then, a directional response can be found by multiplying the magnitude of the intensity vector by the squared magnitude of a microphone polar pattern in the direction of the intensity, for each time-step. The sign of the directional response should be restored from the pressure.

This technique is capable of modelling arbitrary polar patterns, additionally with frequency-dependent characteristics, by processing the intensity and pressure outputs with a set of polar patterns corresponding to responses at different frequency bands, then band-passing with linear-phase filters at each of these bands, and summing the results. This technique could be used to crudely model HRTF receivers, if the HRTF data was approximated by a small number of frequency bands.

The main draw-back of this method is that its accuracy is dependent on the directional error of the underlying mesh, which is much higher for the tetrahedral mesh (a mean of 8%) than other mesh topologies (which generally have a mean of less than 1% directional error).

Choice of Method

It appears that the intensity calculation technique is superior for arbitrary receiver emulation in the DWM. Because no mesh oversampling is required, the simulations will be much cheaper to run than if the Blumlein difference technique was used. Also, the intensity calculation technique can be adapted to any mesh topology, whereas the Blumlein difference technique is most suited to rectilinear meshes. Finally, if a B-format ambisonic output is required, the intensity calculation technique can be used to model receivers with the particular polar patterns required. The intensity calculation technique appears capable of everything that the Blumlein difference technique can do, but more generalised.

The directional accuracy of the tetrahedral mesh is the only possible drawback of the intensity method, and testing will be required to see if this poses any issues in practice.

Implementation and Testing

Initial Testing

First, the magnitude of the directional estimation error for the tetrahedral mesh was calculated, to see whether the magnitude of the error was constant with frequency, or whether it was reduced as the frequency range was reduced. This was achieved using the directional estimation error analysis technique by Hacıhabiboglu et al. (2010). At the maximum angular frequency supported by the DWM ($\pi / 4$), Hacıhabiboglu found a maximum directional estimation error of 19.01 degrees, with a mean of 8.1 degrees.

Results from replicating the analysis were similar, finding a maximum directional estimation error of 15.0 degrees at the maximum angular frequency. The discrepancy in results may be due to the step size used for the analysis. The elevation and azimuth were both incremented in intervals of $\pi/100$, whereas the step size is not mentioned by Hacıhabiboglu.

It was noted that, by reducing the maximum angular frequency, the maximum error could also be reduced. A maximum angular frequency of $\pi/8$ led to a maximum error of 7.5 degrees, while a maximum angular frequency of $\pi/16$ led to a maximum error of just 3.75 degrees. The results from these tests are shown in figure 1 - note the different scales on the colourbars in each plot. It appears as though the maximum error is proportional to the maximum angular frequency simulated in the mesh. Therefore, for applications where high directional accuracy is required, the mesh can simply be oversampled to a point where the error of the desired output bandwidth is within acceptable limits.

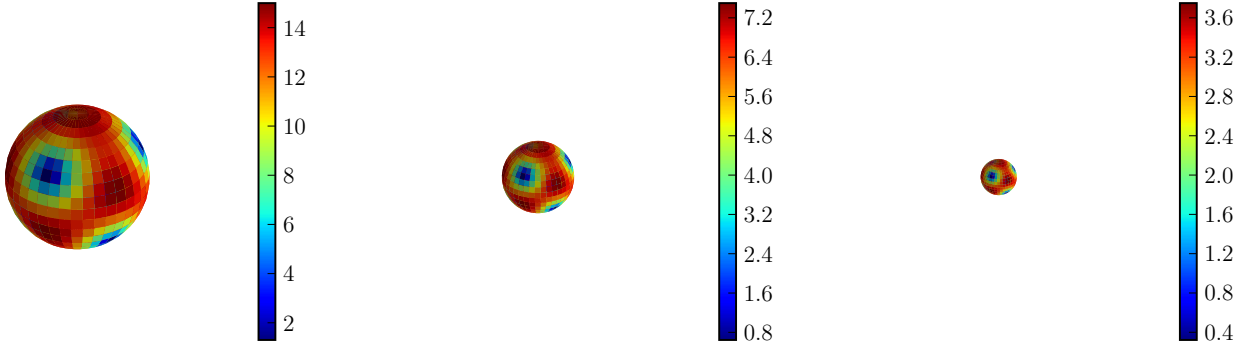


Figure 1: Directional estimation error for tetrahedral mesh, with frequency (from left to right) $\pi/4$, $\pi/8$, and $\pi/16$

Implementation

Next, the intensity calculation algorithm was implemented as part of the tetrahedral waveguide mesh. Here, the tetrahedral mesh is simulated on the GPU, so that each node of the mesh can be updated in parallel.

For each iteration of the waveguide simulation, the differences in pressure between the output node and its immediately adjacent nodes are found, and placed in a vector. This vector is multiplied by a transformation matrix, which is derived from the relative positions of the adjacent nodes, giving a three-element vector. This new vector is multiplied by the negative inverse of the ambient density, giving an approximation for the derivative of velocity with respect to time. Finally, this value is multiplied by the sampling period, and added to the previous velocity, to give the current velocity. Intensity is found by multiplying together the velocity and the intensity.

The output of the waveguide is an array of pressure-intensity pairs. This array can be post-processed to produce the directional receiver output. The same array could be processed using several different receiver patterns to find the directional output for multiple coincident capsules, without needing to re-run the waveguide simulation.

The post-processing step is very simple. For each item in the array, the magnitude of the intensity is found. This value is multiplied by the square of the polar-pattern attenuation in the direction of the intensity vector, and then the square root is taken. Finally, the sign of the pressure is copied to this output value.

If the polar-patterns have multiple frequency bands, this process should be repeated, once for the polar pattern in each band, and then the outputs filtered using linear-phase bandpass filters, and summed together. Optionally, a diffuse-field-response filter can be applied to this summed output.

The author's implementation has two different receiver post-processors - a parametric microphone processor, which can simulate a microphone with a pattern that is smoothly changable from omnidirectional to bidirectional, and an HRTF processor, which uses an 8-band simplification of HRTF responses to produce multiband outputs which are then filtered and summed. The architecture of the program is such that additional receiver processors would be trivial to add.

Testing Procedure

Hacıhabiboglu et al. (2010) describes a testing procedure in which a simulated directional microphone is placed at the centre of a room measuring 4.1 x 5 x 2.1 m with a cubic mesh topology. The mesh spacing used is 1.35cm,

facilitating a sampling rate of 44.1kHz, and a theoretical upper mesh frequency limit of 11.025kHz. The walls and floor have slightly different absorption coefficients, and are modelled as phase-inverting 1-D boundaries. The mesh is sequentially excited using Gaussian pulses with a variance of four samples, positioned in a circle with 1m radius around the microphone. The waveguide simulation is run for 200 steps, then the energies of the resulting responses are calculated and normalized.

The testing procedure for the author’s implementation is similar. However, the mesh topology used is tetrahedral, and boundaries are modelled as nodes with a constant zero pressure. Boundary modelling in the mesh as a future research topic for the author, at time of writing.

Three different microphone polar patterns are simulated: omnidirectional, bidirectional, and cardioid. The output of the waveguide is split into octave bands, starting at 80Hz and ending at 10.24kHz. Then, the energies in each band are graphed for each direction. This method was chosen as the waveguide mesh generally has increased directional estimation error and dispersion error at higher frequencies. The results are shown in figures 2, 3, and 4.

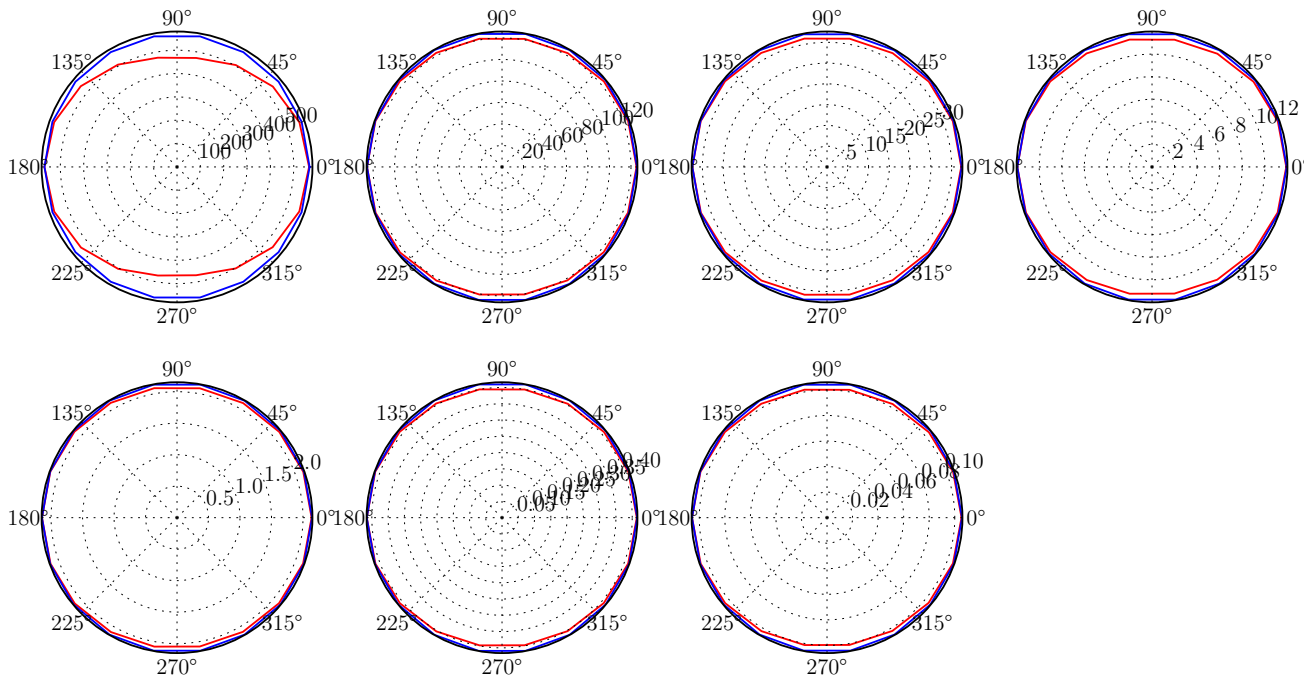


Figure 2: Energy at a simulated omnidirectional receiver. Red is the measured energy, blue is the desired energy. Frequency increases in octaves from left to right, top to bottom, starting at 80Hz.

It can be seen from the results that the modelling of directional receivers definitely works to some extent. This being said, the different polar patterns can be clearly identified and differentiated. Somewhat strangely, the shapes of the polar patterns appear consistent over the frequency range, with the greatest error in the lowest frequency band, while a greater degree of error might be expected in the higher frequency bands. This low-frequency error is most evident in the omnidirectional plots. It appears as though this technique will be suitable for modelling of directional sources, although the low-frequency performance is troubling, especially as the desired application of the waveguide mesh is to low-frequency modelling. The solution to this issue is probably oversampling the mesh.

Bibliography

Hacıhabiboglu, H., Günel, B., & Cvetkovic, Z. (2010). Simulation of directional microphones in digital waveguide mesh-based models of room acoustics. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(2).

Southern, A., & Murphy, D. (2007a, October). Methods for 2nd order spherical harmonic spatial encoding in digital waveguide mesh virtual acoustic simulations. *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*.

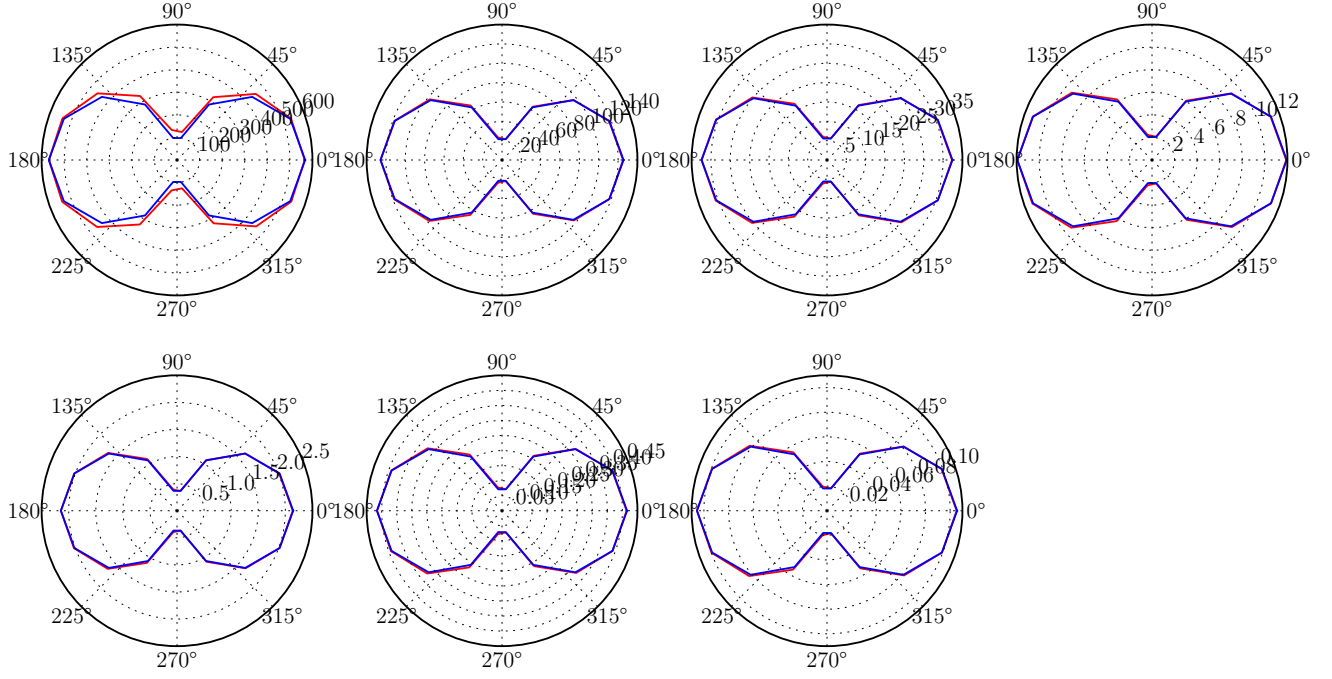


Figure 3: Energy at a simulated bidirectional receiver. Red is the measured energy, blue is the desired energy. Frequency increases in octaves from left to right, top to bottom, starting at 80Hz.

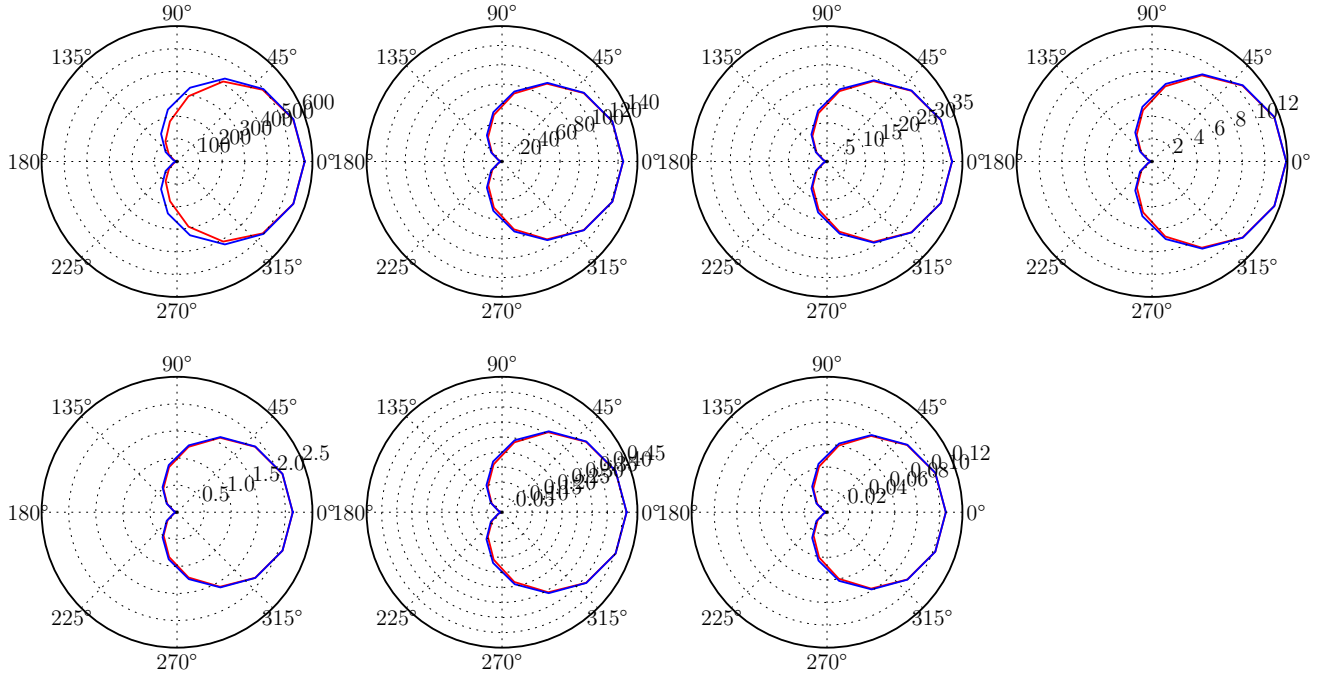


Figure 4: Energy at a simulated cardioid receiver. Red is the measured energy, blue is the desired energy. Frequency increases in octaves from left to right, top to bottom, starting at 80Hz.

Southern, A., & Murphy, D. (2007b, September). 2nd order spherical harmonic spatial encoding of digital waveguide mesh room acoustic models. Proceedings of the 10th International Conference on Digital Audio Effects.