The Audio Spotlight:  
An Alternative Approach

P. Hong, IMTC, Georgia Institute of Technology

Abstract

The purpose of this project was to design a system that would actively direct audio in desired directions and locations. Although there already exist proven methods of achieving this goal, including ultrasound modulation in [8] and sound direction by parabolic dishes, a digital signal processing (DSP) and antenna theory approach was developed for this system.
1 Introduction

The applications for a directive sound system range greatly from public address systems to personal notification systems in a room. Because of the many potential applications, system development becomes a higher priority than before.

Here, the system to be developed consists of a computer equipped with an 8 audio output soundcard and the amplifiers and loudspeakers associated with each audio channel. The computer uses a multi-track audio program to playback individual sound files on each channel. These sound files were generated by MATLAB with delays and gains dependent on the loudspeaker positions as well as the desired optimal observation point coordinates within a specified room which can also correspond to a direction depending on the equations used in MATLAB and whether a near-array or far-array approach is desired.

This system differs greatly from the aforementioned systems for directing sound. To briefly explain:

- The ultrasound modulation system simply does that, it modulates a carrier wave that operates in the ultrasound frequency range with the desired audio signal. By doing this, the system takes advantage of the non-linearities of air at those frequencies thereby causing demodulation of the desired signal for hearing. This scheme provides very high directivity as noted in [8]. The primary disadvantage is the damage it causes to the human auditory system although this is currently being studied and will hopefully be circumvented in the future.

- The parabolic dish approach essentially is an antenna approach. Just as in antenna applications, a relatively omni-directional loudspeaker is placed at the focal point of a parabolic dish, pointing towards it. Consequently, when a desired sound is sent to the loudspeaker, it acts like a point source letting out a spherical wave to reflect off of the dish creating a directive beam of sound. This system has already been implemented by several companies. Its primary disadvantage is its bulk; the parabolic dish has to be relatively large in order to accommodate the longer wavelengths of lower frequencies and to provide a suitably sized listening area.

Both systems also require mechanical means to change the location they are “illuminating” with sound. The system we wish to develop will not require mechanical means for steering and will not be particularly bulky either. Steering is controlled by the delays of the signal to each loudspeaker, i.e. the steering is electrically controlled. A diagram of the set up can be seen in Figure 1: Diagram of System Set Up.
Figure 1: Diagram of System Set Up
2 Theory

The ideas behind the system are relatively simple. Each signal from each speaker is delayed such that they will all arrive at the desired location or direction at the same time. When applied correctly to a linear array of loudspeakers, these signals will add up in the area in front of the array. This theoretically works because there will be perfect constructive interference in the desired location/direction and destructive interference elsewhere. The equation for calculating the delays in the near field is:

\[ t = (d - r) / c \]

where \( t \) is the delay, \( d \) is the longest distance of the array elements to the observation point, \( r \) is the distance of the element in question, and \( c \) is the speed of sound. For far field, the delay is equal to the dot product of the speaker position with the unit direction vector divided by the speed of sound all subtracting the minimum delay of all the elements, i.e.:

\[ t = s \cdot r / c - tm \]

where \( s \) is the element position, \( r \) is the unit vector in the observation direction, \( c \) is the speed of sound, and \( tm \) is the minimum of all the delays of all the elements.

Although this obviously does not provide perfect directive control, i.e. there is sound outside of the desired area, but this beamforming technique is a proven method used in radar and array signal processing as shown in [4]. To suppress the inevitable side-lobes, the gains can be set to Hamming of Hanning window coefficients to provide tapering to the linear array of loudspeakers. Although this increases the main lobe’s beamwidth, the side-lobe suppression is typically worth it.

Another effect for consideration is that of side-lobes due to element spacing. The loudspeakers need to be placed less than half the shortest wavelength accommodated by the array. Theoretically, there are at least two views on this.

- From antenna theory, this can be justified simply because greater spacing between elements leads to higher sidelobes as seen in Chapter 7 of [7].

- From a DSP point of view, the elements can be thought of as spatially sampling the area. Then, from the Nyquist criterion, the above limitation seems reasonable as noted in [4].
3 Simulation

All simulation was done in MATLAB primarily based on [5] which describes a computer simulation program for loudspeakers and loudspeaker arrays. This was relatively simple to code in MATLAB and provided a means for estimating the 2D and 3D directivity functions for all the array configurations. An example of visual output is shown in Figure 2: Directivity Plots.

Additionally, after referring to [7], another program was developed that generated the sound pressure graph seen in Figure 3: Sound Pressure Plot. These simulations did not take reverberation into account, which proved to be quite a problem in the latter stages.
A listing of the MATLAB programs is included on disk, and a brief description of the functions follows in

Table 1: Program Listing.

<table>
<thead>
<tr>
<th>File Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mysim.m</td>
<td>General script file that has exercises the important functions and plots the directivities and sound pressures.</td>
</tr>
<tr>
<td>directivity.m</td>
<td>Generates a directivity matrix and theta and psi vectors.</td>
</tr>
<tr>
<td>pressures.m</td>
<td>Generates a sound pressure matrix and the y and z coordinates.</td>
</tr>
<tr>
<td>getpressures.m</td>
<td>A function that calculates the sound pressure from a given array and observation point.</td>
</tr>
<tr>
<td>fardelays.m</td>
<td>Generates the delays for a given array in the far field.</td>
</tr>
<tr>
<td>neardelays.m</td>
<td>Generates the delays for a given array in the near field.</td>
</tr>
<tr>
<td>tones.m</td>
<td>Generates .wav files that contain delayed sinusoids.</td>
</tr>
<tr>
<td>tones2.m</td>
<td>Generates .wav files that contain 2 delayed sinusoids (can have different delays) of different frequencies.</td>
</tr>
<tr>
<td>hear.m</td>
<td>Generates .wav files of delayed input sounds.</td>
</tr>
<tr>
<td>linearpos.m</td>
<td>Creates a vector of positions for array elements in a line.</td>
</tr>
<tr>
<td>circlepos.m</td>
<td>Creates a vector of positions for array elements lying on a circle.</td>
</tr>
<tr>
<td>circularpos.m</td>
<td>Same as circlepos.m except it takes inter-element spacing as an argument.</td>
</tr>
</tbody>
</table>
Table 1: Program Listing
4 Implementation

In order to optimize sound in a particular location, different time delays and gains are used on the different signals sent to the different speakers. The signals, as .wav files, were produced or pre-processed in MATLAB, which easily computed the delays and applied the gains. For the delays, because we are using digital systems, we simply rounded them to an integer number of samples for delaying. This is reasonable because the sampling rate used was relatively high (22050 Hz). So the new equation for our digital delays becomes:

\[ dd = \text{round}(d * FS) \]

where \( dd \) is the digital delay, \( d \) is the actual delay, \( FS \) is the sampling frequency, and the round function simply rounds the product, in this case, to the nearest integer.

whose basis is found in [6]. The loudspeakers were placed in a line all facing the same direction, side by side against each other. Consequently, the “point source” equivalents (the center of the woofer) were spaced 13.51 cm apart from each other. Both Hanning (since Hamming windowing would have effectively reduced the number of elements to six) and unity gains were used during simulation and playback. Additionally, the typical frequency used was 1200 Hz because of the Nyquist criterion mentioned earlier as well as the sharp directivity functions generated at this frequency during simulation.

Once the .wav files were created, they were loaded into Cool Edit Pro, a multi-track software package.

From there, the signals were played on a Gina soundcard (an 8 analog output soundcard). Then each output channel was amplified individually (using two Denon POA-8300 3 channel amplifiers and one Denon AVC-2800 amplifier for the remaining two channels).

An additional test to mention was to create locations of complete silence by delaying half of the loudspeaker signals by half a wavelength. Although the simulations reported good results, very well-defined valleys of low sound pressure, actual implementation did not fare as well in the original lab. Although valleys were present, few were in calculated positions and those that were in their correct positions were not always minima. The final test in a less reverberant environment did create the expected results.
5 Testing and Data

The different tests employed for this project included audio inspection as well as measurements taken with a standard microphone and oscilloscope. Although the audio inspection method is not necessarily objective, it served its purpose with the applications in mind.

To test the audio equipment, an oscilloscope was used in conjunction with a microphone. Initially, all the outputs of the Gina soundcard were tested to make sure there was no phase delay between output channels; this was confirmed. Next, the output of each channel was measured against the Gina soundcard output to measure the delay across the audio equipment. The results follow in Table 2: Measured Delays. These measurements were taken directly off of the oscilloscope and no attempt was made find correlations based on frequency to find an overall delay figure because that is unnecessary. The purpose here is to confirm that at given frequencies, there is a relatively uniform delay for each channel. These numbers indicate acceptably close channel delays.

<table>
<thead>
<tr>
<th>Channel</th>
<th>200 Hz</th>
<th>400 Hz</th>
<th>600 Hz</th>
<th>800 Hz</th>
<th>1000 Hz</th>
<th>1200 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.80 ms</td>
<td>2.50 ms</td>
<td>1.74 ms</td>
<td>1.38 ms</td>
<td>0.12 ms</td>
<td>0.10 ms</td>
</tr>
<tr>
<td>2</td>
<td>3.10 ms</td>
<td>2.50 ms</td>
<td>1.75 ms</td>
<td>1.32 ms</td>
<td>0.12 ms</td>
<td>0.10 ms</td>
</tr>
<tr>
<td>3</td>
<td>3.40 ms</td>
<td>2.50 ms</td>
<td>1.75 ms</td>
<td>1.38 ms</td>
<td>0.12 ms</td>
<td>0.11 ms</td>
</tr>
<tr>
<td>4</td>
<td>3.30 ms</td>
<td>2.50 ms</td>
<td>1.75 ms</td>
<td>1.38 ms</td>
<td>0.11 ms</td>
<td>0.12 ms</td>
</tr>
<tr>
<td>5</td>
<td>3.15 ms</td>
<td>2.50 ms</td>
<td>1.75 ms</td>
<td>1.38 ms</td>
<td>0.11 ms</td>
<td>0.11 ms</td>
</tr>
<tr>
<td>6</td>
<td>3.20 ms</td>
<td>2.50 ms</td>
<td>1.75 ms</td>
<td>1.38 ms</td>
<td>0.12 ms</td>
<td>0.11 ms</td>
</tr>
<tr>
<td>7</td>
<td>3.20 ms</td>
<td>2.50 ms</td>
<td>1.75 ms</td>
<td>1.35 ms</td>
<td>0.11 ms</td>
<td>0.13 ms</td>
</tr>
<tr>
<td>8</td>
<td>3.90 ms</td>
<td>2.50 ms</td>
<td>1.75 ms</td>
<td>1.40 ms</td>
<td>0.13 ms</td>
<td>0.10 ms</td>
</tr>
</tbody>
</table>

Table 2: Measured Delays
6 Conclusions

Although the project did not succeed, there are several issues to be gained. The simulation software, after extensive analysis appears to be functioning correctly. This was verified by several individuals, and the test patterns run on the software match the known data. Unfortunately, the software needs to predict reflections from hard surfaces in order to be more useful for most indoor situations.

On a similar theme, the testing environment was less than optimal. Although there was a lot of space, plenty for the equipment and to walk around and try to distinguish peaks and valleys in sound level, there were too many hard surfaces providing reflections of sound. Consequently, many of the results are suspect because of this echoic environment. Initially, we had thought that the 1/r attenuation would prevent echoes from contributing too much, but after consultation with different professors, we came to the conclusion that the echoes were distorting our results. We sought an anechoic chamber for testing, however, to no avail.

We were able to test in a relatively reflection-free area on the roof of a parking structure. The only reflecting surface was the cement floor, but its contribution was minimal. This environment provided an excellent testing area as all the simulation results generally were confirmed. Although we did not have a sound meter available, audio inspection was possible and appropriate from the intended applications.

Another possible source of discord is the audio equipment. We lack the equipment to successfully determine the transfer function of the equipment used to play the signals. Although we did use an oscilloscope to measure the delay at several frequencies, there could be coupling between loudspeaker elements, thereby creating oddly delayed signals or signals pointed in unexpected directions.

Although we gathered the expected results from the roof excursion, they did not fare well enough. The simulations had predicted high directivity as well as promising pressure levels for attenuation outside of the desired listening area. However, these levels were not distinct enough for actual applications probably due to the human ear’s logarithmic listening attributes. Reducing sound levels by attenuation never brought the sound pressures to the point that the sound was confined distinctly to that area. Additionally, the constraint that we be in a non-reverberant environment limited the initial plethora of applications.

For future directions, considerations to make include:

- Increasing the number of elements. Although this gets costly for broad-band applications since there would need to be a different array (with different delays and gains from each other) for different frequency bands, if a simple delay control mechanism was found, this approach is feasible.
Also, a hybrid system of some kind incorporating already-existing techniques may be the most promising technique. This is imperative for a broad-band system because inter-element spacing at high frequencies is unattainable with available transducers.
References

The Author

Paul Hong was born in the fine state of Delaware but lived out most of his early life in Bloomfield Township, Michigan. He received his B.S.E.E. from the University of Michigan, Ann Arbor, in 1997 and currently is pursuing his Ph.D. in electrical engineering from the Georgia Institute of Technology, Atlanta, Georgia under Professor Mark Smith. Paul was employed by the Interactive Media Technology Center of the Georgia Institute of Technology as a research assistant for his work on this project although his interests are mainly concerned with digital signal processing, more specifically, wavelets and filter banks. He is a member of IEEE.