Wave-Field Synthesis: Conceptual Understanding and Practical Application

Ben Collins

University of Michigan
Sound Recording 331
December 1, 2006

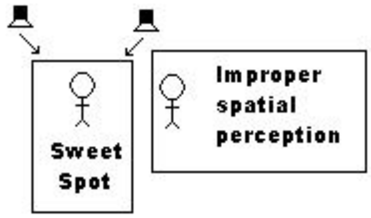
Audio recording is, by all traditional standards, an extremely new form of art. We have been recording and processing sound for less than a century, and as such, the procedure itself is constantly being revised and improved upon. New techniques are always being developed to more accurately replicate not only the timbral quality of a sound, but it's spatial orientation and its environment.

Naturally, audio recordings began as simple mono signals. Separate sounds were recorded and mixed together, but due to the single loudspeaker, a listener was forced to perceive all audio as coming from the same direction.

Stereo was the next logical breakthrough in sound recording. Suddenly, an engineer was able to place multiple mono signals into a stereo field to give the perception of distance between instruments or voices. In conjunction with volume and reverb, listeners were suddenly presented with a two dimensional spatial orientation.

Surround sound, in terms of emulating a sonic environment, was the next major breakthrough. Listeners were still presented with a two-dimensional sound field, but were additionally given the perception of audio in front and back.

These techniques work quite well for conventional audio recording. In order to localize a mono signal, they rely on the volume and timing differences between speakers. Unfortunately, the aforementioned techniques intrinsically assume that the listener is sitting in a position equidistant from all loudspeakers. If a listener moves away from this "sweet spot," the illusion of space is damaged or lost entirely.



This becomes a major problem when dealing with large listening audiences. Everyone is presented with a different representation of spatial separation. When virtual sound sources are placed on a one-dimensional region in space, you might think of it as a perfect horizontal line floating in the air; a person standing perpendicular to the line would be able to see its full length, but anyone at a different angle would perceive it as being shorter.

To provide a larger listening environment, a new technique has been developed known as Wave-Field Synthesis (WFS), first conceived by Berkhout in 1988 [1]. WFS has been designed to overcome the innate problems associated with traditional stereo recording. This paper will investigate the theory behind WFS, and its potential application for practical use.

Overview

WFS overcomes the challenges of traditional stereo by utilizing an array of smaller speakers to emulate a particular environment. In short, it's based on the idea that a single audio signal can be accurately represented by a sum of smaller, secondary signals at an arbitrary uniform distance from the source. We begin to see a more intelligent method of spatial localization based on emulating a continuous distribution of secondary monopoles. In a perfect situation, we would have an infinite number of speakers with an infinitely small amount of space between them. Engineers quickly realized that it's possible (thankfully) to sufficiently emulate this scenario with a smaller number of speakers [2].

They devised an array of small loudspeakers that act to re-create the secondary monopole response to a primary audio signal. The result is a recreation of the exact conditions under which the sound file was recorded. Additionally, the effective listening area becomes much larger. We can now utilize a large physical area in which one can listen to a sound and perceive it as originally intended by the artist.

With the advent of virtual sound sources by means of WFS combined with environmental modeling techniques, we can simulate a much more realistic environment than has ever been possible before. The practical potential for this seems almost limitless.

Huygens Principal:

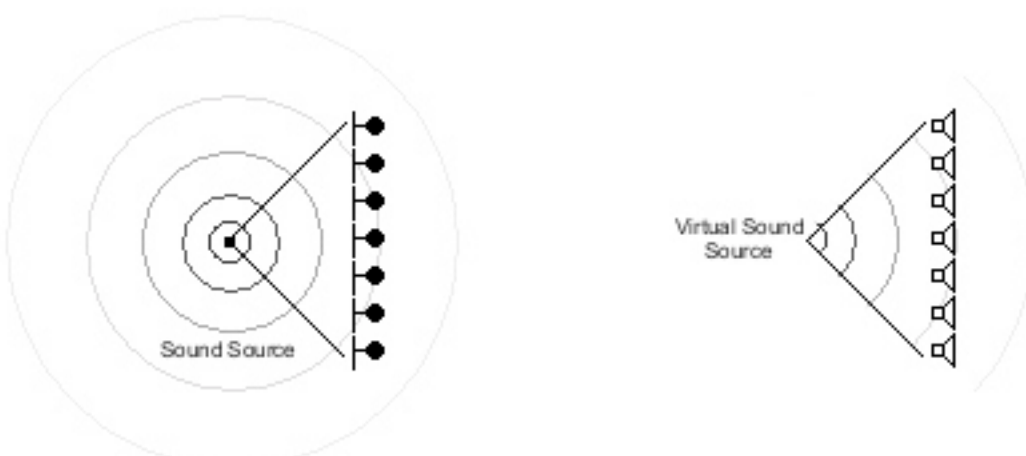
In order to better understand the nature of WFS, it's important to know the Huygens principal [5]. This is the idea that in an isotropic, homogenous medium, a particle that oscillates will transfer its motion to particles around it. This continues equally in all directions, creating a spherical wave.

We can begin to see the importance of this principle with regard to sound. A single point source of sound will vibrate the air around itself, and a wave will propagate in all directions. The state of oscillation at an arbitrary distance to the source can be determined by summing up all the individual secondary oscillations around the sphere.

We have previously discussed the issues associated with traditional recording. It's apparent that stereo doesn't accurately recreate sound sources. A mono signal panned between two speakers is a quick and simple way to recreate a directional sound, but if a listener walks around, it's quickly obvious that the effect is manufactured.

Imagine that we were to place a loudspeaker in a room for each individual sound we wanted to play back. There would be a center speaker for vocals, speakers for each individual drum sound, and a speaker for each instrument. We could theoretically place these in such a way that they perfectly emulate a performing ensemble. A listener could walk back and forth in the listening field while maintaining perception of space.

Clearly, if an engineer could have that much control, there would be no problem with spatial perception. But this isn't the case. We can, however, simulate these sound sources by means of the Huygens principle. To our ears, there is absolutely no difference



between hearing a direct sound source and hearing the effect it has on secondary monopoles. By simulating the secondary reaction of distant air particles, we can more accurately localize a virtual point of sound. [2]

Reverberation*:*

It's important to note that by placing virtual sound sources, we are not automatically recreating an environment. To accomplish that, we must also simulate the natural reflections of the primary sound source in a room.

This is commonly done using external reverb processing or other artificial forms of reverb. Another common type of reverb simulation, which is important in context with WFS, is convolution reverb. A dry signal is convoluted with a room's recorded impulse response, and the result is a more natural sounding reverb that accurately simulates a true space.

Spatial Rendering through Wave Field Synthesis

Now that we have discussed the Huygens Principle and convolution reverb, we can begin to understand the method by which WFS is implemented.

By placing an adequate number of speakers close together, we can emulate secondary air vibrations and give the perception of a virtual sound source. In order to do this, an array of microphones must be placed in a line near the origin of sound. The secondary vibrations will be picked up by the microphones, and can be played back through an array of similarly arranged loudspeakers.

A new effect that we can now create is that of a plane wave. This creates the impression of a sound source being focused in the same direction across the entire listening area. Plane waves in two dimensions are similar to sin waves in one dimension, with lines being formed perpendicular to the direction of propagation [7]. This gives rise to more creative options than were previously available.

Even more exciting is the potential for projecting virtual sound sources into the listening field itself. By means of WFS, one is able to simulate a sound source as coming from a spot *within* the listening environment. It would be possible for a listener to move around in the sound field and experience the same effect that they would if there were a loudspeaker playing right next to them. We are suddenly presented with an audio analog of a hologram, known as a "holophony [2]."

We can also update our method for recreating room reflections. Naturally, it's still possible to use traditional reverb techniques alongside WFS. Running dry signals through digital or analog reverb processors still creates the illusion of room reflection. But using the fundamentals of sound field synthesis, we can recreate the impulse of a room much more accurately.

Convolution reverb is already a much more realistic method of applying a particular room's characteristics to a sound. But rather than using traditional methods for measuring the impulse of a room, we can now utilize WFS concepts to better capture the environment. By setting up perpendicular microphone arrays, we will have better directional information with regard to a room's impulse. In addition to the artistic implications, we can use this to obtain a more accurate auralization of potential future halls. [1]

Virtual Panning Spots:

Due to the nature of the recording techniques necessary to produce WFS, it becomes more difficult to properly capture a large ensemble such as an orchestra. As a result, sound engineers have found a way to combine traditional stereo microphone techniques with WFS by means of “virtual panning spots (VPS).”

A VPS is a virtual loudspeaker that represents a particular point in space [2]. An engineer would set up, for instance, five microphones in a wide arrangement across an ensemble using more traditional miking techniques. As a result, we are given the ability to reproduce an image of the group in five channels. It is now possible to stretch or compress the image of the ensemble as appropriate for the recording, and use this arrangement for our WFS processing.

Software Application:

As with any recording process, there is an innate wall between mathematical theory and practical creative application. In order to bridge this gap, interface software has been developed to allow an artist to focus on the creative side of a project without worrying about writing their own algorithms for WFS.

WONDER [3] (Wave field synthesis Of New Dimensions of Electronic music in Real-time), by Marije Baalman, is one such software interface. WONDER allows a user to compose the individual movements of a sound source within a piece, or to simply specify a grid of points to use. This software has been tested by a number of composers with a high degree of success.

WONDER is just one example of an interface program for use with WFS. Undoubtedly, as the method becomes standardized, we will see more and more software (and hardware) designed to elegantly control sound localization.

Advantages and Disadvantages to WFS:

One can clearly see the advantages of WFS over traditional stereo or surround techniques. The listening environment is much larger, the perception of sound is much more natural, and an artist is given much more creative control than was previously available. Additionally, it has been shown that speech is much more intelligible when localized through WFS. Videoconferencing would benefit from this idea.

WFS is also innately better than binaural processing, which utilizes head-related transfer function to create the perception of spatial localization. Binaural techniques require a user to be wearing headphones, which immediately makes group listening very difficult.

There are several problems with our current techniques for sound field synthesis, however. In a perfect environment, the speaker array would be infinitely long. Since all practical speaker arrays are *not* infinite, we must deal with edge effects. The reflections from the edge of the speaker array will be perceived as after echoes. Engineers are still wrestling with a proper method for fixing this. [2]

Additionally, it's impossible to create an *exact* replica of the sound field without an infinite number of infinitely small speakers that are infinitely close together. Until we figure out how to do that, we'll be stuck with the issue of discreteness. That is to say there is spatial aliasing that occurs due to the distance between loudspeakers. For frequencies that are above F_{alias} (the spatial aliasing frequency), the time difference between two speakers will interfere with proper perception. [2]

There are also issues associated with the listening environment. Unless the listeners are in an anechoic chamber, there will be natural reflections of the room that may interfere or even take over the perception of synthesized space. There is investigation being done on algorithms that would compensate for particular listening rooms. [2]

Applications and Possibilities for the Future:

Sound field synthesis is very new, and we have barely scratched the surface of different possibilities. The potential is limited only by our own creativity. But some ideas for practical usage may include:

Cinema: Movie theaters provide a good environment for the speaker arrays necessary to do WFS. Surround sound is good, but the ability to localize virtual sound sources and project them into the listening area would conjure up new possibilities for film sound [4].

Teleconferencing: Since speech intelligibility is an innate problem associated with video conferencing, WFS would allow speech to have a more focused source.

CARROUSO: The European project CARROUSO combines WFS technology with MPEG-4 video compression. The project showed that it's possible to accurately capture and transfer sound field information from one room to another. CARROUSO is the first project to combine video broadcasting with immersive audio simulation. [6]

Re-Processing Stereo: It has been shown that a stereo mix, when reproduced as two plane waves at 30°, provides a much better image by means of WFS. [4]

3D Entertainment: In conjunction with hologram technology, a new field of entertainment has opened up. We can already project a 3D image into the audience area. Combining this with a sound source focused on the same location would allow us a vast array of new possibilities. There could be virtual concerts and movies, or even Star Wars style communication devices. [4]

Conclusion:

WFS is a much more sophisticated and elegant method of sound field replication than any we have seen before. Its disadvantages are relatively small compared to the new possibilities that will arise. We have yet to see the full range of practical applications that this technology makes possible.

References:

[1] Boone, Marinus and de Vries, Diemer, 1999, “Wave Field Synthesis and Analysis using Array Technology,” *Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York, Oct. 17-20, 1999*

[2] Theile, Gunther, 2004, “Wave Field Synthesis – A Promising Spatial Audio Rendering Concept,” *Proc. Of the 7th Int. Conference on Digital Audio Effects (DAFx’04), Naples, Italy, October 5-8, 2004*

[3] Baalman, M.A. J., 2004, “Application of Wave Field Synthesis in Electronic Music and Sound Installations,” *Linux Audio Converence 2004*

[4] Boone, Marinus, 2001, “Acoustic Rendering with Wave Field Synthesis,” *ACM SIGGRAPH and EUROGRAPHICS campfire, Snowbird, Utah, May 26-29, 2001*

[5] H. Geiger, and K. Scheel, “Handbuch der Physic,” band VIII. Verlag Julius Springer, Berlin, 1926, Cited in [2]

[6] CARROUSO project, http://www.idmt.fraunhofer.de/projects/carrouso, last visited on November 29th, 2006

[7] David Raymond, 2006, “Plane Waves,” http://physics.nmt.edu/~raymond/classes/ph13xbook/node22.html, New Mexico Tech, last visited November 28th, 2006

Illustrations by Ben Collins, with the aid of Microsoft Paint.