Sonic Presence™, Imagery of the Mind

Russell O. Hamm – 8/30/2018

ABSTRACT

50 years of refining traditional stereo microphone techniques have not produced sound recordings that are even close to realizing the goal of creating a live experience. Cognitive psychology suggests that this short coming is due to the near absence of the Head Related Transfer Function (HRTF) in stereo recordings. Sonic Presence is a new recording method using acoustic sensors that you wear on your head. The HRTF is embedded while you make a recording. The result is a three-dimensional sound image that sounds as if you are there.

1 Introduction

Sonic Presence is a new method for sound recording that captures sound with three-dimensional realism. We developed Sonic Presence by taking a fresh, new approach to sound recording. We re-examined our sense of hearing with its fantastic ability to locate sounds coming from all around us. Then, we asked the question: Why when we listen to a media file does the sound stay trapped inside our head?

Research in the field of cognitive psychology suggests the Head Related Transfer Function (HRTF) and its associated time and level differences are critical for cueing our mind's auditory perception. Yet this function is mostly absent in today's sound recordings. Traditional microphones, and the complicated techniques for using them, do not adequately capture the HRTF. We've replaced them with our Spatial Microphone, a small, highly sensitive device that you wear on your head. It embeds the HRTF into a file while you are making a recording.

When you listen to a Sonic Presence file your mind detects the embedded spatial cues. The sound image expands outside your head and beyond. Left, right, in front, and behind – you hear the full 360-degree soundstage all around you. Sonic Presence captures these spatial cues the way your mind has evolved to process them. Instead of trying to create an audio image with signal processing, our Spatial Microphone captures sound with embedded spatial cues that let your mind create the audio image. The result is incredibly realistic; it's as if you are there.

2 Sound localization

Over one hundred years ago, Lord Rayleigh [1] in his treatise "Duplex Theory of Sound Localization," described the basic principles of how we hear sounds coming from different directions. There are two principles, hence the word "duplex." The first principle is time difference. When a sound originates from a source directly to our right, we'll hear it first with our right ear and then, a fraction of a second later, with our left ear. The time difference is miniscule. About 600 millionths of a second, but our mind can detect it. The scientific terminology for this time difference is: Interaural Time Difference (ITD).

The second principle is level difference. Because there's a head located between our ears, the sound coming from the source on our right will be louder in the right ear and softer when it arrives at the left ear. That's because our head blocks the sound creating a level difference or shadow. This level difference is not so simple to understand as the time difference. Our head is spherically shaped. Its interaction with sound waves creates level differences that are frequency dependent and quite complex. Our mind is amazingly sensitive to these level differences. The scientific terminology is: Interaural Level Difference (ILD).

There is a vast body of mathematical literature written during the last century analyzing the interaction of waves with rigid bodies. We've studied these works to gain an understanding of how sound waves behave when they encounter a spherical object, specifically our head. Sound waves create pressure zones as they impinge upon and pass around our head. These effects influence our mind's sense of direction, creating a sense of spaciousness and presence.

3 Current Stereo techniques

Stereo sound recordings became widespread in the late 1950s, although their invention dates to the 1930s. In concept, stereo recording is supposed to capture sounds and reproduce them in a way that recreates the live experience. We have two ears, so the theory behind stereo says that two channels of sound should perform this illusion.

Sound engineers have many options for positioning traditional microphones to make stereo sound recordings. Stanley Lipshitz [2] describes many of the options in his paper "Stereo Microphone Techniques". In summary, the combination of microphone spacing and directional angles produce variations in the ITD and ILD. The basis for all these stereo recording techniques is Rayleigh's 100-year-old Duplex Theory of Sound Localization.

Binaural introduces effects of the human head into the stereo recording process. Microphones are inserted into the ears, positioned as close as physically possible to the eardrums. Since most humans find these intrusions into their ears uncomfortable, binaural recording has not proven to be popular. Francis Rumsey [3] summarizes recent research in binaural in his report titled, "Whose head is it anyway?". New developments in binaural playback [4]

The fact is, 50 years of refining traditional stereo microphone techniques have not produced sound recordings that are even close to realizing the goal of creating a live experience. Clearly, something is missing.

What's missing is the spherical shaped object in the middle: our head.

4 Sonic Presence method

Sonic Presence method for sound recording addresses the many flaws of traditional microphone techniques and binaural. We've replaced traditional microphones with our Spatial Microphone, a small, highly sensitive device that you wear. It embeds the HRTF into a file while you are making a recording, so it captures the sound the way you hear it.

Our Spatial Microphone uses two acoustic pressure transducers. They're attached to ear hooks made of a malleable material, so you can shape them to fit your ears. Figure 1. In contrast to binaural, you position the transducers in front of your ears near the temples. The shape of human heads is much more uniform in this area, making the HRTF similar for different individuals. By moving the transducers in front of the ears, we minimize the sonic effects of the pinnae whose shape differs widely between individuals. Moving the transducers forward also reduces the recording angle, which enhances the center image and fills in the hole in the middle, which is a chronic binaural problem. Our transducers are not inserted into your ears like binaural, so there is no ear canal resonance or physical discomfort. You can enjoy the sound while you are making a recording.

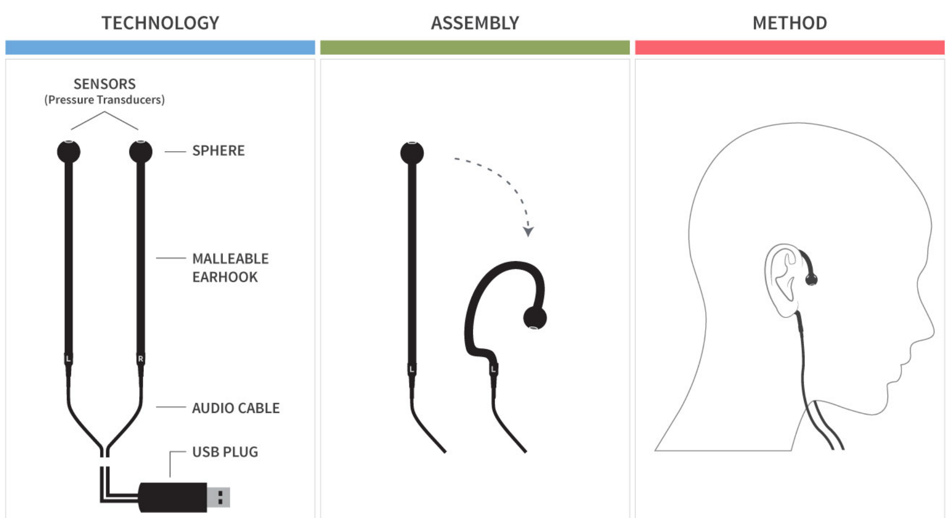

Figure 1– Sonic's Spatial Microphone uses two pressure transducers mounted on each side of your head. They're attached to ear hooks made of a malleable material. You shape them to fit your ears. The transducers embed the HRTF as you make a recording.

Traditional microphone techniques do not capture the dynamic effects that our head has on sounds. Our head is not exactly a sphere, it's an ellipsoid, but it behaves in a way like a sphere. We'll show how these spherical effects give our mind important spatial cues that we use to construct a three-dimensional sound image.

4.1 Bright spots

When a direct sound wave collides with a sphere it creates a pressure zone called a "bright spot". There is a buildup in sound pressure at the bright spot because the rigid surface of the sphere reflects the sound wave back onto itself. Figure 2. The behavior of the bright spot is a function of the wavelength of the sound relative to the diameter of our head. For this discussion, we'll use a diameter of 175mm for the human head. That's in agreement with much of the technical literature published over the past 100 years.

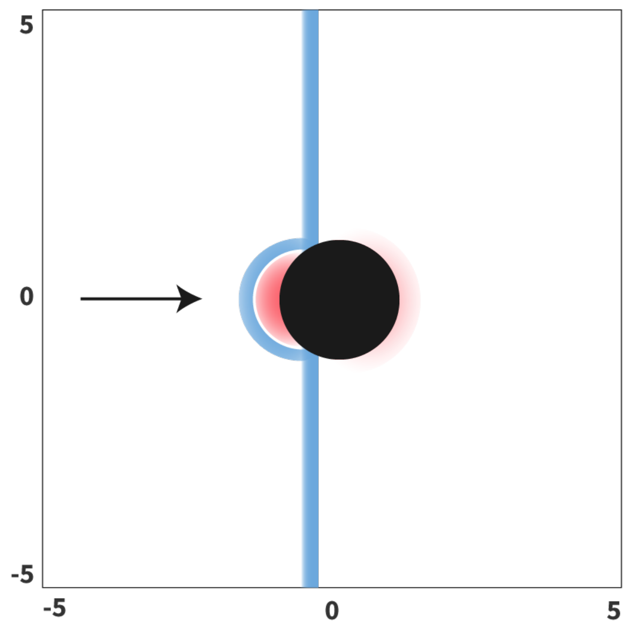

Figure 2 - A direct sound wave colliding with a sphere creates a pressure zone called a "bright spot". There is a buildup in sound pressure at the bright spot caused by the rigid surface of the sphere reflecting the sound wave back onto itself [5].

Muller, Black and Dunn [6,7] of Bell Labs describe the behavior of a sound wave interacting with a sphere in their paper published in 1938. They express the pressure behavior at the sphere's surface using a ratio of diameter to wavelength. Wavelength relates directly to the speed of sound and is inversely proportional to frequency. At long wavelengths, frequencies below 100Hz, the head has no effect. There is no bright spot. Sounds coming from all directions produce equal pressure on the sphere's surface. At a wavelength of two-tenths the head's diameter, which is a frequency of 250Hz, a bright spot starts to form. There is a gradual increase in pressure at the bright spot as the frequency increases over the next three octaves. When the frequency reaches

2,000Hz, a wavelength equal to the head's diameter, the sound pressure at the bright spot nearly doubles (an increase of ~6dB). Thereafter the pressure in the bright spot remains constant as the frequency increases out to the limit of our hearing.

4.2 Direct vs. diffuse sound

Direct sounds create pressure zones or bright spots where the wave impinges on our head. Sounds coming from many random directions are called "diffuse" sounds. These are sounds caused by the reflections of direct sounds off surfaces like the walls, floor and ceiling of a room. They are scattered in many directions and their amplitude diminishes as they collide with our head. Direct sounds by contrast produce bright spots that increase their amplitude. This spherical effect of our head helps our mind to differentiate direct sounds from diffuse sounds. Effectively, our head creates a natural "presence boost" for direct sounds by increasing their amplitude. We call this the Natural Presence Boost of the Human Head. Figure 3.

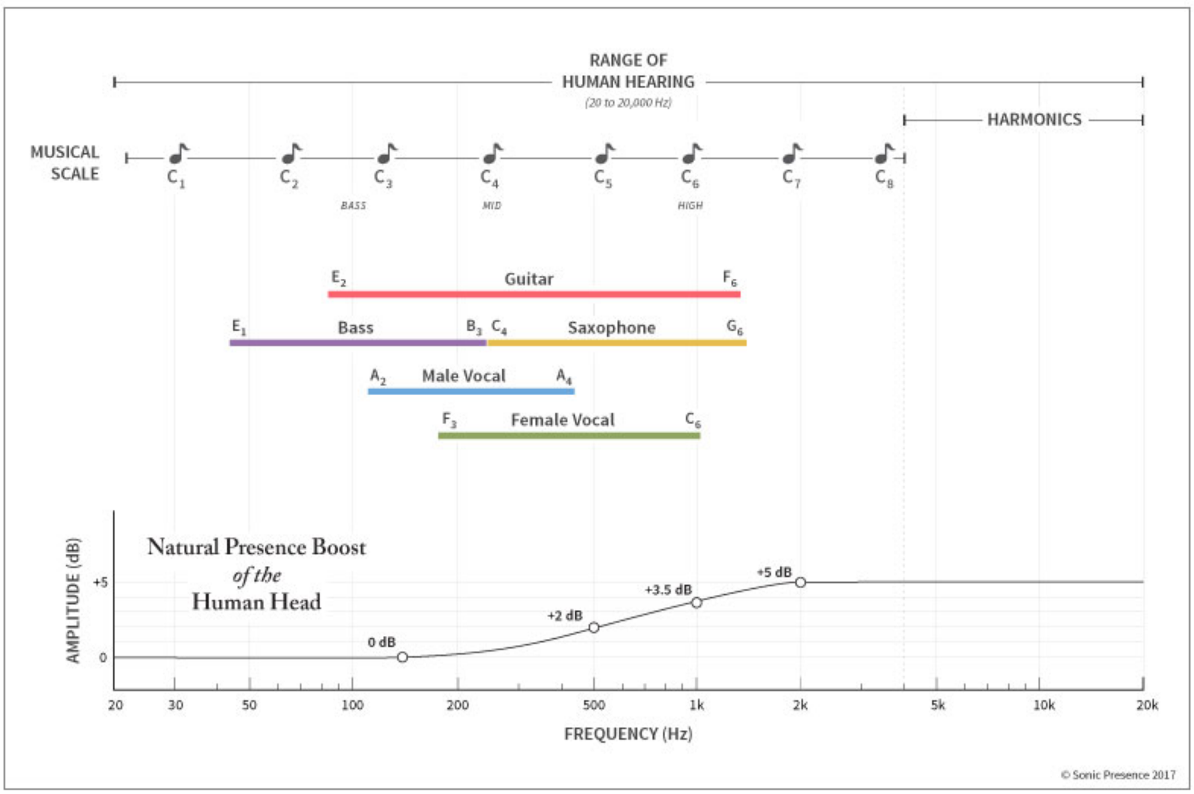

Figure 3 - The Natural Presence Boost of the Human Head begins at a frequency close to Middle C in the center of the male vocal range. It reaches its maximum at around C_7 (2,093Hz) just above the highest note of the soprano voice. Musically, the notes in this three-octave vocal range are the most important ones for melody.

When we hear performers at a concert, they sound close-up. They have a natural “presence.” This is the spherical effect of our head emphasizing the direct sound over the diffuse. When we make a traditional stereo recording this effect of our head is missing. Our mind does not get the spatial cues it needs to differentiate the direct sound from diffuse sounds. Without these cues our mind perceives all sounds as diffuse and that’s why they sound far away.

4.3 Natural presence boost

We can make some interesting musical observations about the Natural Presence Boost of the Human Head. The range of human hearing spans the frequency range of 20Hz to 20,000Hz. Within this range lie the notes of the musical scale from C_1 at 33Hz to C_8 at 4000Hz. Sounds at frequencies above the musical scale are mostly harmonics, which define the character of sounds we hear. Vocal sounds and most common musical instrument sounds lie in the middle five octaves. Bass notes range from a mid-sounding B_3 (246.94) extending downward to a very low sounding E_1 (41.2Hz). A guitar’s range of notes begins an octave above the Bass and extends upward over four octaves to its highest note F_6 (1396.91Hz). Middle C, or C_4 (261.63Hz), is in the middle of the male vocal range, while the female vocal range lies an octave higher. What’s interesting about the Natural Presence Boost of the Human Head is that its effect begins at a frequency which is close to Middle C, right in the center of the male vocal range. It reaches its maximum at around C_7 (2,093Hz) just above the highest notes of the soprano voice. Musically, these three-octaves are the most important ones for melody. It’s not surprising then, that these same three octaves are where our spherically shaped head naturally enhances the presence of the sounds we hear.

Listening tests, which we will be publishing in a companion paper, show Sonic Presence recordings improves our ability to correctly locate the direction of sounds by 300% when compared to traditional microphone techniques.

4.4 Directionality

Technically, our acoustic transducers are omnidirectional. They are equally sensitive to sounds coming from all directions. At long wavelengths, bass frequencies, our head has no directional effect. We hear bass sounds equally in all directions. But, because of the interaction of sound waves with our head, the transducers become directional at higher frequencies. This change in the transducer’s directionality from omnidirectional to unidirectional results from a combination of the bright spots and the head’s shadowing effect. The bright spot creates an increase of amplitude pointing in the direction of the sound’s origin, while the sphere’s shadow causes a decrease in amplitude in the opposite direction. This differential in amplitude causes a cardioid like directional characteristic. As the direct sound increases in frequency, the beam narrows, becoming even more directional at higher frequencies.

4.5 USB Interface

Sonic's Spatial Microphone converts variations in sound pressure into an electrical signal. Using digital signal processing, we format the signal to USB standard, stereo audio with a resolution of up to 96kHz, 24 bit. You can record this format on most mobile devices that you carry in your pocket including iPhone and Android. The three-dimensional sound of our Spatial Microphone also perfectly complements the 4K video images captured by today's best mobile devices.

5 Conclusion

Sonic Presence replaces traditional microphones with our Spatial Microphone, a small, highly sensitive device that you wear. It embeds the HRTF into a file while you are making a recording, so it captures sound the way you hear it.

When you listen to a Sonic Presence file your mind detects the embedded spatial cues. The sound image expands outside your head – you hear the full 360-degree soundstage all around you. Sonic Presence captures these spatial cues the way your mind has evolved to process them. Instead of trying to create an audio image with signal processing, our Spatial Microphone captures sound with embedded spatial cues that let your mind create the audio image.

Our Spatial Microphone uses two pressure transducers mounted on your head in front of the ears on the temples. Positioning the transducers away from the ears and moving them forward solves several problems that plague the binaural recording method. Sonic Presence captures a well-defined stereo sound stage with a center image that is clearly positioned in front of the listener. In contrast to binaural, there is no "Hole in the Middle". Soloists stand out with special "presence" because of the bright spots created by direct sounds interacting with our head. The Natural Presence Boost of the Human Head and the HRTF are embedded in the recording. Sonic Presence sounds live.

Acknowledgements

Special thanks to Bob Ludwig for asking provocative questions about our Sonic Presence recordings that inspired me to write this paper, and to Chris Van Deusen for the illustrations.

References

[1] Rayleigh, "Theory of Sound", Vol 2 MacMillan and Company, 1929

[2] Lipshitz, S., "Stereo Microphone Techniques", AES #2261 – 1985 May

[3] Rumsey, F., “Whose head is it anyway?” JAES Vol 59, 9 - 2011 September.
[4] Choueiri, E. “Optimized Crosstalk Cancellation for 3D Audio over Two Loudspeakers”, https://www.princeton.edu/3D3A/Publications/Choueiri_3D3A_OptimalXTC.html
[5] Feuillade, C., “Animations for visualizing and teaching acoustic impulse scattering from spheres”, JASA Vol. 115,5 – May 2004.

[6] Muller, Black and Davis, “The Diffraction Produced by Cylindrical and Cubical Obstacles and by Circular and Square Plates”, JASA Vol 10,6 - July 1938.

[7] Wiener, F. M. – “Sound Diffraction by Rigid Spheres and Circular Cylinders”, JASA 19,3 – May 1947

Further Reading Material

Blauert, J, “Spatial Hearing”, MIT Press, Cambridge, MA.- 1997
Bore and Peus, “Microphones”, Georg Neumann Berlin, 1999
Levitin, Daniel – “This is Your Brain on Music”, Penguin Books 2007
Olsen H, “Elements of Acoustical Engineering” D. Van Nostrand Co., 1940