YOUR HOME THEATRE

It’s All In The *Acoustic* Details

Part II

provided by Widescreen Review
The Essential Home Theatre Resource™

Norman Varney

It really is all the small details in the acoustics, which make for a believable, powerful experience *over and above* the equipment. It doesn’t matter whether you have a $50 system or a $500,000 one—if noise and sound-quality issues are not controlled, your experience will be low res. Don’t bother buying new equipment in order to get better performance until you are sure you have done all that you can with your setup and environment first, to allow it to perform optimally. I’ve seen systems with a $50,000 loudspeaker stuffed into a corner. How great can that sound? Certainly not like it is supposed to. Imaging and timbre will be a mess. People pay for accuracy and then destroy it without knowing or understanding what they’re doing. Sure, some don’t care, but most don’t know what they’re missing.

In Part I we discussed how noise control is the foundation for sound quality, and its correlation to resolution and dynamics. We also brought up how most people have never had a good demonstration and don’t know what a good experience is. While listening to music on a good system (system meaning the entire electro-acoustical system, which includes the room and its entities), you can experience what seems like transportation to where the musicians are playing. You can realize the size of the space and where each musician is located within it. You can feel the emotion of the player’s soul as he plays his instrument, and it can cause tears of joy or sorrow to well up in your eyes. A vocalist’s breath heard here or a string heard touching a fret board there, makes the experience more personal and intimate. Subtle clues are not missed in movies, and they become much more suspenseful and real when small details are heard and imaging is accurate. Music videos are more fun when tonality is true. You become familiar with, and appreciate, the sound characteristics of The Royal Albert Hall, Steinway pianos, a Neumann U87 microphone, and Black Diamond guitar strings, etc.

In Part II, we will continue to use more video calibration terms to relate to audio, and delve further into sound quality—what it means to us, and how to achieve it. I will also vent more pet peeves, or “rants” of mine. But before we start, we need a little introduction as to how we hear and compute sound. It should help you understand and appreciate the complexity of sound quality, and how important “system” setup and calibration is to its performance capabilities, and our end experience.

When we describe light and sound in basic terms, we talk of waves. Respectively, color and pitch are frequencies, brightness and loudness are intensities. When describing these energies, we must also include time. Frequency and intensity are a little easier to understand than time. For sound, time is our primary measurement for localization, along with frequency, intensity, and our aural memory. We quickly compute all of this data together to come to conclusions, which stimulate reactions.

We may incorporate several different “tests” in order to determine what we see and hear. With monocular vision we see a world that has breadth and height, but not depth. We have about 140 degrees of vertical vision, limited by our brows and cheeks, and about 150 degrees of horizontal vision (see Figures 1 and 2). Imagine a straight road that goes on for miles, with telephone poles running along its shoulder. With only one eye, we can see that the poles on the side of the road near us appear far apart, and that they become narrow and almost

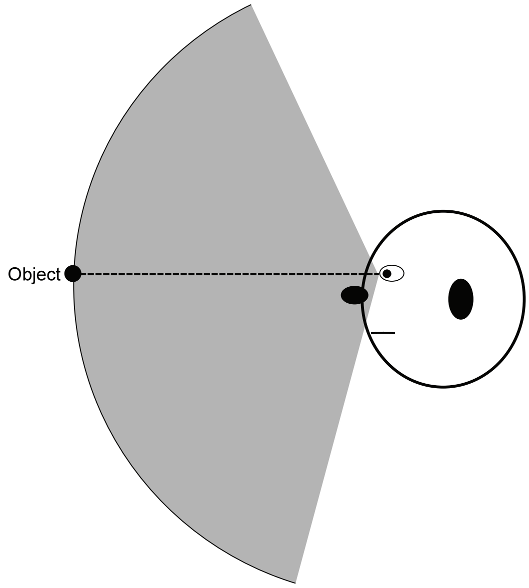

Figure 1. Our approximately 140 degrees vertical field of view.

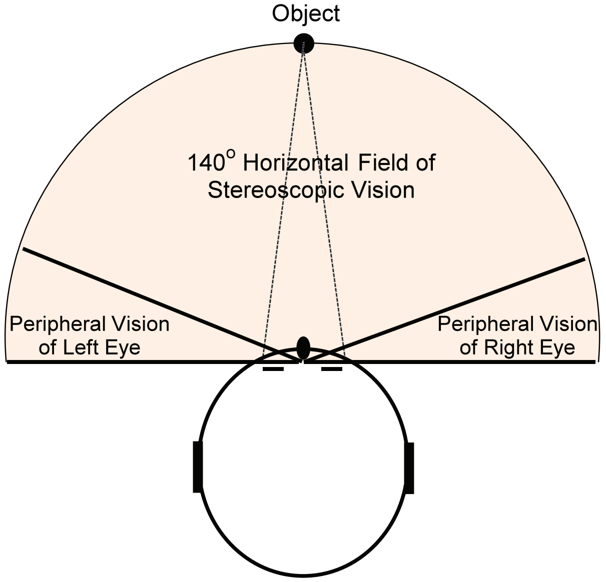

Figure 2. Each eye has a range of about 150 degrees of horizontal vision. Together we have about 180 degrees of horizontal view. Where they overlap is where we have our binocular field of view.

touch each other in the distance. We also can see that the telephone poles gradually look smaller and smaller the further away they are. We can guess the distance of the approaching car by relating the size to the surrounding objects in our view. There are no visual depth cues telling us about distance, only experience pulled from our visual memory.

Again with one eye, imagine looking out a window at the sky and you see a dark object against the blue sky. At first you can't figure out what it is. Is it a tiny bug on the window just a few feet way or is it a big airplane miles away? You move your head to view it from a different angle in order to relate its *position* to objects in the foreground and background. Instantly its identification, size and distance are no longer in question. With binocular vision, we have an overlap of horizontal vision and achieve approximately 180° of view. With this stereoscopic vision our pair of eyes allows us to focus upon a single object from two slightly different angles. The visual cortex combines these two pictures to form roundness and solidity to objects, and provides our primary means of perceiving depth in the world.

Depth Experiment: *Try capping a pen with one eye closed while someone else holds the pen. Somewhat similarly, try closing your eyes while listening to the world around you, then fold or cover your ears to hinder them just a little and notice the spaciousness collapse.*

We know that the world stretches all around us, but we can't see behind our head and must rely on our ears to draw our attention to what is not in our view. Imagine the sound of someone calling out to you from a hilltop, versus someone calling from across the room. From the hilltop, we can hear from the strain in their voice that they are shouting. This clue tells us that they must be far enough away that they think they must shout in order to be heard. We also detect that the voice is not as loud as it would be if that person were shouting near to us. We observe that the sound quickly diminishes and that there are no reflections, so the sound must be originating from outside. We then try to fix on its location by seeking other clues as to what happens to the sound as it approaches us. We note timbre characteristics of high- and low-frequency loss and the overall low-loudness level.

If sound originates somewhere in front of us, we have the ability to pinpoint its horizontal position within 1 degrees of arc. Our vertical resolve is not nearly as accurate. In fact, we don't hear very well directly behind and below us (see Figures 3 and 4), and become confounded with locating sounds directly overhead. This is due to the shape and position of our pinna (external part of the ear). Spatial cues are resolved by the disparity of sound arriving at our two ears and the fact that our head interferes with them. Generally speaking, frequencies below about 700 Hz are largely localized by phase (time) differences, and above 1.6 kHz are largely defined by intensity differences. The mid-range area of overlap offers our keenest degree of resolution because both methods of deciphering are at work. Another timing method of localization is called the *law of the first wavefront*. This has to do with our brain fixating on the sound heard by the ear nearest the source first. In the horizontal plane, we are able to resolve time arrivals between 630 μs and 28 ms (after which they are heard as discrete echos). For example, if centered between a typical stereophonic loudspeaker system, while listening to a monaural recording, the sound appears to come from a phantom loudspeaker between the two because equal energy is emanating from both loudspeakers. If you move just one inch closer (about 0.10 ms) to one of the loudspeakers, the image will shift from center towards the closer loudspeaker. If you move just one foot closer (about 1.0 ms), the image will seem to come from only the closer loudspeaker. Again, in the median plane we don't resolve as well. Why?

1. **We have only one ear on each side of our head.** If we had two ears on each side of our head, one facing forward and one facing the rear, we could hear equally to the rear, as to the front.

2. **Both are aimed forward.** If we could direct the pinna around like many animals can, we could locate sounds better behind us. If we had one ear slightly lower than the other (like owls), we would be good at locating sound distances above and below us.

3. **They are about 7 inches apart.** Physically perfect for our primarily mid-range communicating world due to their wavelengths. In addition, our head casts an acoustical shadow on the ear furthest from the source. This helps us correlate the differences impinging on both ears by making differences of intensity, frequency, and time larger than if they were, say, on top of our heads.

After sound travels through a space, likely altered by being absorbed and reflected by various obstructions, it enters the pinna

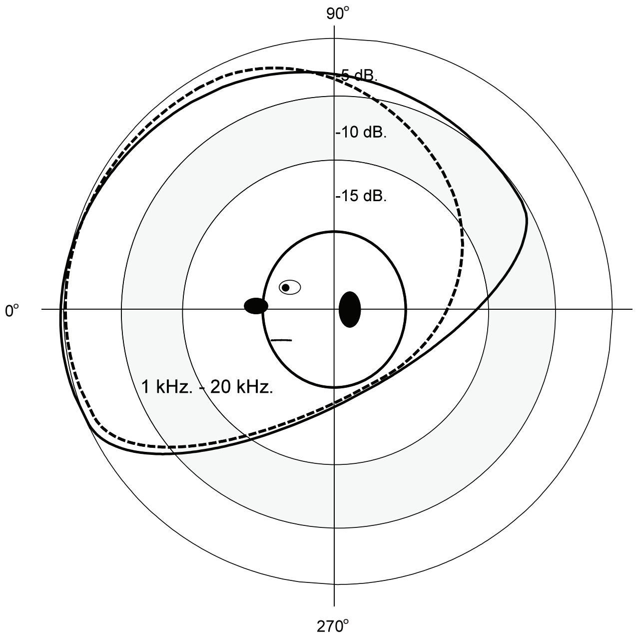

Figure 3. Approximation of how we hear speech in our vertical field.

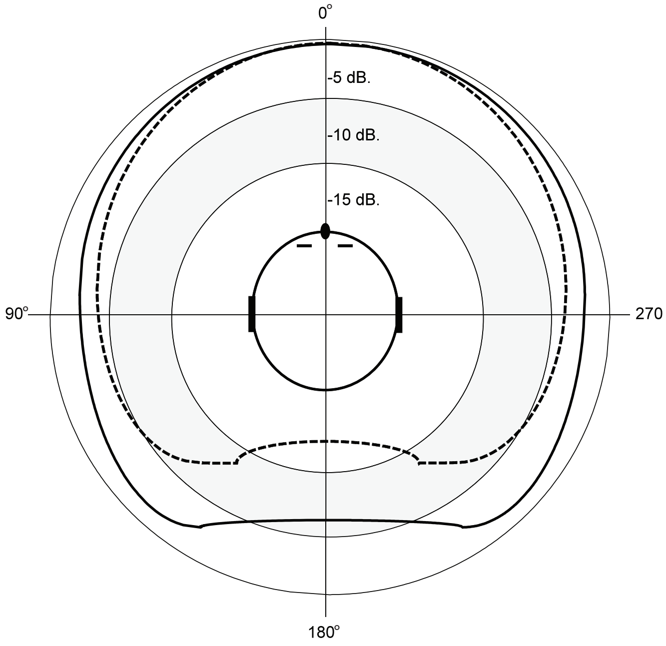

Figure 4. Approximation of how we hear speech in our horizontal field.

and travels into the ear canal, onto the eardrum and ossicles, to the cochlea, where cochlear fluid meets the nervous system. Here, in the *organ of Corti*, hair cells are triggered to convert vibrations of hydraulic pressure into electrical impulses. Our brain then compares the differences of signal input of both ears. This is performed by nuclei called *olivary bodies* in the brain stem. These signals then go to the cerebral cortex for more processing, where the information is deciphered and related by different parts of the brain, to trigger memory, emotion, and reaction.

Rant: *The idea of averaging multi-channel sound for all listeners is not a good one.*

First of all, you can't average time, reverberation, or energy at more than one location, and why would you want to average the frequency response of all the seats? Each seat will still be unique. The point in the room where time, energy, and frequency converges should be the center of the listening space. Any other seat will be compromised to a degree, regardless. Why compromise every seat? As depicted in Figures 5 and 6, more seats will exhibit better sound without averaging frequency response. In addition, audio nirvana can be experienced in the primary seat.

Rant: *The primary seat is not centered between the loudspeakers.*

I can't tell you how often I see this done these days—dedicated, otherwise symmetrical rooms, where there is no seat offered in the middle of the auditory scene. The best sound is had by straddling an armrest or two (see Figure 7). This was not done in past centuries.

In Part I we made some comparisons of video technology with audio, as they relate to noise control. Let's continue with some relating to sound quality.

Gray Scale (frequency response)

In video, gray scale has to do with colors (frequency) being linear from dark to bright. If white is not actually white, but is slightly blue, all colors will be slightly tinted blue. In audio we want the same linear frequency response from each channel. For example, if we position a loudspeaker or ourselves in a corner, the soundtrack will sound bass heavy.

Color (loudness)

In video, color level has to do with intensity, same in audio. Again, we want all frequencies reproduced with the same intensity, or loudness, in each channel. For example, if you take a seat closer to one loudspeaker than the others, you will experience errors in the "audio picture" and miss out on information intended by the artists.

Tint (time alignment)

In video, this has to do with phase (signal timing relative to another) affecting color hue or tint. In audio, time alignment affects timbre (tonality) and spatial cues (imaging). When signals arrive at your ears, the brain deciphers their tonal characteristics, as well as the disparity between the direct sound and the reflected sounds, and also the arrival times to both ears. A signal's attacks, decays, and harmonic structure allow us to recognize the difference between a B sharp struck on a piano and B sharp blown through an oboe. The distinction between the direct signal vs. reflected signals allows us to recognize whether the instrument is being played in a large hall or a small room. The inequality of arrival time to each ear allows us to decide where the instrument is located in relation to our head position.

Uniformity (symmetry)

In video, this is about being able to project the same color and light output across the entire screen. For audio, this can relate to having everything set up electro-acoustically the same on the left side of your head, as the right. For example, if your seat, your loudspeakers, and/or your room is set up asymmetrically, wave fronts, both direct and reflected, will arrive at your ears asymmetrically, causing time, energy, and frequency distortions. This results in timbre and image distortions. Add more channels and you add more possibilities for error.

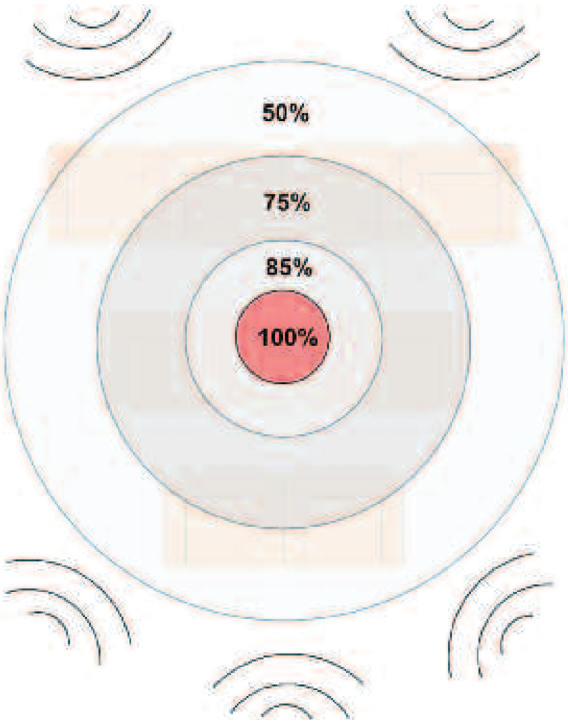

Figure 5. Approximation of performance percentage of a properly set up and calibrated 5.1 system at different seats.

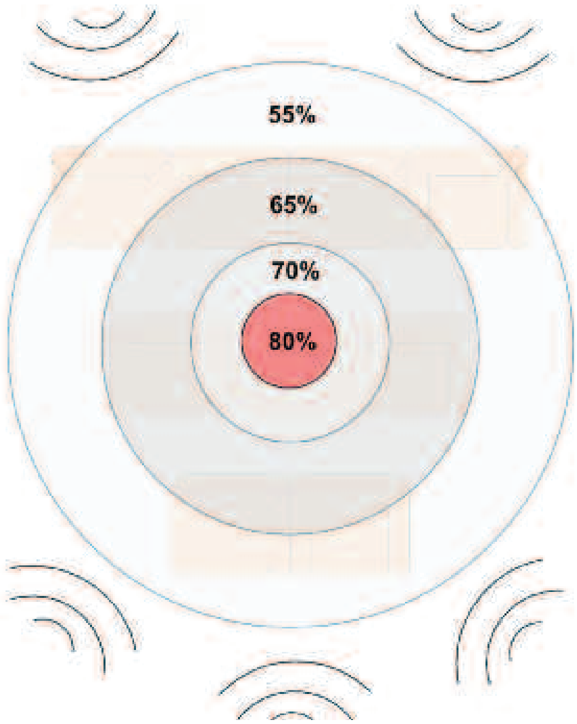

Figure 6. Approximation of performance percentage of a 5.1 system, which has been frequency "averaged" or "diluted" for all seats.

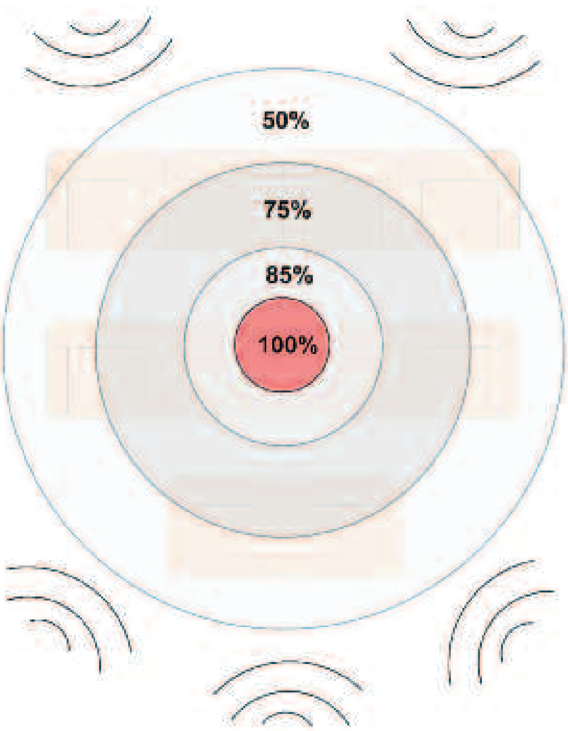

Figure 7. Approximation of performance percentage for the common "no middle seat" scenario.

Aspect Ratio (senses, emotional responses)

This one is maybe a bit far-fetched, but I'm going for it anyway. In video, a selection of aspect ratio results in two observable outcomes: 1) physical size of the picture, and 2) picture content information. In audio, image size has a lot to do with frequency range. For example, a small bookshelf size loudspeaker cannot throw a life-size sound image of many instruments or sound effects. Imagine listening to Tchaikovsky's "1812 Overture" on a boom box; everything would sound miniaturized.

Many years ago I conducted a series of subjective tests, concluding that big sound with a small picture made the movie experience more involving than did a big picture with small sound. This is true for many reasons, the most obvious is that big sound means broader frequency and dynamic range, which better mimics real life. But another important factor is the tactile information that is sensed because of it. With the additional sense added to the mix, a whole new set of clues can be added to the scene. There are some movies that my 11-year-old daughter prefers to watch on our small system rather than on our big dedicated one because they're too scary! Take for instance the difference in sound between watching *Harry Potter And The Order Of The Phoenix* on a TV versus a full-blown system. Things are felt physically. Not only low-frequency rumbles, but sudden loud noises that cause you to jump. The sounds and the soundfields are truer to life because there is more information to make them match more closely to our life experiences. Now imagine listening to Mussorgsky's "Night On Bald Mountain" on a small system; you can't feel the weight of the tympani or the tickling of the string section, nor can you feel the emotion of the piece, if the dynamics and frequency range is on a small scale. You'll miss the meaning of the notes played at *fff*, and you will miss the whisper-quiet *ppp* notes entirely.

I hope that this little introduction to sound quality has helped you to better understand and appreciate our amazing ear/brain mechanism, and how important good setup is to achieving good performance. In Part III we will cover low-frequency sound quality issues such as: room shape, size and dimensions, room modes, door locations, resonance damping, subwoofer locations, etc., along with more rants. **WSR**

About The Author

Norman Varney is the owner of A/V RoomService, Ltd, an acoustical firm specializing in sound quality, noise control, power quality, and HVAC, offering design, modeling, testing and voicing services, and many acoustical products. Prior to A/V RoomService, Norman was with Owens Corning at the Science & Technology Center, where he was a Senior Engineer with the Acoustic Systems Business as the Acoustic Design Center Lead. Prior to Owens Corning, Mr. Varney worked at Music Interface Technologies, where he designed critical listening and viewing environments, AC line conditioners, and video cables and was Director of the Custom Installation and Home Theater divisions. He was the lead for the development of the 2C3D and 5C3D Certification programs, which recommended structural, electrical, and system component setup parameters for Spectral, Avalon, ASC, and MIT. While there, he designed the very innovative and elaborate electrical system for the Scoring Stage of Lucas Film's Skywalker Ranch. Mr. Varney has written many articles for numerous magazines over the years, as well as given seminars and participated on panel discussions regarding acoustics. He became a member of AES in 1981 and has contributed to the Characterization and Measurement of Diffusion Coefficient Standards, and the Recording Academy's Producers & Engineers Wing Recommendations for Surround Sound Production. He continues to study and develop the science of subjective acoustic value-to-performance relationships. Norman can be contacted at www.avroomservice.com, or normanvarney@avroomservice.com.