



The Art & Science Of Surround Sound

Time For A New Surround Sound Standard—Part II

GARY REBER

Introduction

In this continuing series of articles related to advocating a new surround sound standard, the following letters to the editor provide an excellent platform to address this subject once again.

My position is simply that we should set as the goal a new surround sound standard of mixing and playback of imaging-specific 360-degree soundfields or what I term three-dimensional “holosonic” spherical surround” soundfields. Once the goal has been agreed on then the solution(s) will be eminently apparent.

Surround Sound Standard Reasoning

Dear Gary:

I applaud your engagement with the industry to move the state-of-the-art forward. However, I missed several topics in your discussion of this issue which I think would be of interest to your readers. First I would like to see a discussion of the psycho-acoustic research behind your recommendations. Second, while you put forward your position forcefully, you did not outline the position taken by THX® and their reasons for taking that position, although, by reading between the lines we can get some idea of their position.

On the first point, it is clear to me that pinpoint imaging on the sides probably doesn't work too well, as our ears aren't in the right location for it. What does the research show? How does good stereo imaging on the sides compare to the “fuzzy wuzzy” soundfield when a person is faced 90 degrees away? Maybe you are right, but you should present the research that proves it. If this research hasn't been done, maybe you can foster it at a local university. Additionally, it is also clear that if the LCR loudspeakers are pretty close together, you don't need to rely on the stereo imaging effect to position instruments and voices across the sound stage; rather they are hard positioned L, C, and R. Since the stereo imaging collapses as you move away from the sweet spot, this hard localizing to an

array of loudspeakers across the sound stage can replace it for creating a larger “semi-sweet” spot for a larger audience. Again, what does the research show? How do humans react with stereo imaging between the center loudspeaker and one of the L or R loudspeakers, versus stereo imaging between the L and R loudspeakers when you are in the sweet spot? Does it really work with the center loudspeaker/L speaker? Again, what does the research show? All of this is very relevant to the discussion, and we enthusiasts would like to know!

Keep up the good work. It's appreciated.

Gene Banman President & CEO,
NetContinuum, Inc., Santa Clara, California,
gene@netcontinuum.com

Editor-In-Chief Gary Reber Comments: I do not agree with your assertion that “...pinpoint imaging on the sides probably doesn't work too well as our ears aren't in the right location for it.” Sidehead perception of phantom imaging is good enough to perceive relative spatial lateral and depth imaging. I was made even more acutely aware of this phenomena when experimenting with the Smyth Virtual Surround™ headphone system, which exactly recreates the sound of loudspeakers positioned in a surround sound system (see the On Screen department in this issue). This realization is important in creating holosonic three-dimensional soundfield presentations in which all channels to some degree, more or less, are acoustically related. Empirical experience has demonstrated that humans can discern imaging laterally with depth perception between two loudspeakers, recreating a good stereo recording when the listener is positioned in the middle between the two loudspeakers. This is not rocket science and you can easily demonstrate this human phenomena to yourself simply by turning your head 90 degrees to your left or right side while listening to stereo recordings with good lateral and depth imaging. This human phenomenon was the basis for the quadraphonic experience with the idea that stereo on four walls could be created using

four equally capable full-range loudspeakers positioned equidistant (and at equal height) at equal 90-degree included angles relative to the sweet spot listening position. In that experience, stereo heard from the same sweet spot was expanded three-dimensionally to surround the listener. Unfortunately, the delivery to the consumer through a two-channel pathway was a complete failure because the necessary matrix encoding and decoding systems of the day were incapable of sustaining the necessary sense of discrete separation between the four channels that were created in the studio environment. In that environment, recording engineers mixed to four discrete analog channels using multi-track tape recorders. Another factor contributing to the human ability to discern phantom listening from the left or right side of our head is the reflective nature of sound waves in a listening room, which propagated all around our head.

Recordings mixed with hard positioned left, center, and right channels will produce imaging when the loudspeakers are closely spaced and the included angle relative to the listener is narrow. Even without the center channel the phantom content in the mix derived from the left and right channels will be heard as intended when listening equidistant to the left and right loudspeakers at a relatively narrow included angle. However, when this “narrow” mix is spread out in playback over loudspeakers positioned at 90-degree or wider included angles relative to the listener, phantom imaging begins to reveal “holes” in the mix between phantom center and hard left and hard right. That is why I have always advocated that recording engineers mix at wide included angles to insure seamless lateral and depth imaging in the soundstage created by two loudspeakers reproducing stereo. By doing this the imaging will hold up even when the included angle is narrower while at the same time supporting wider included angles. This was the basis of mixing in quad.

The presence of a dedicated center channel also allows the surround mixing engineer to create phantom images between the left and



center and right and center loudspeakers, thus further enhancing localization. But because the distance between those loudspeakers are shorter, these "in-between" phantom images are not as strong but still helpful in spot placement of sounds across the front soundstage.

If you have either or both SA-CD and DVD-Audio playback capability, there are scores of recordings that either use the center channel predominately, partially as a center imaging lock, or not at all (relying on a phantom center image). In any of these methods, listening from the crosshair sweet spot will provide a strong, definitive center image. Phantom imaging degrades both in stereo and in quad-based holosonic soundfield setups when the listener is not listening from the crosshair sweet spot. Even with a discrete hard center channel to anchor the center image, if the listener is seated to the left or right of center the lateral phantom imaging will shift in prominence to the left or right loudspeaker's output depending on the listener's position relative to his or her proximity to the left or right loudspeakers. This is true in the depth sense as well. If the listener is seated behind the sweet spot and farther back closer to the surround loudspeakers, the soundstage created by those two loudspeakers will be predominant depending on the energy placed in that soundstage perspective and the listener's position relative to his or her proximity to the left and right loudspeakers. Similarly, this will be true for sidewall soundstage predominance. This is why I have always advocated the idea of an equidistant, equal height, and equiangular listening position relative to the four main loudspeakers making up left front, right front, left surround, and right surround. The center front and center back (if used in a Surround EX™ or DTS-ES™ configuration) should also be equidistant relative to the listening position and the other loudspeakers. As with the center front, the center back should be reproduced by a single loudspeaker so that good phantom imaging can be assured between the left surround or right surround loudspeaker and the center back loudspeaker. When using two loudspeakers for the center back channel, the potential for good imaging is destroyed.

Generally, it is possible, though difficult, to create a larger "sweet spot" that supports more than one person's optimum listening experience in larger spaces, but not in the typical home theatre, which is a room in a family home. I have argued repeatedly over the past twelve years, since founding Widescreen Review that the focus of home theatre reproduction should be on a typical family household, which consists of a couple married with children. In this household, you will rarely find both adults, or for that matter any of the children critical listeners. That is why the

concept of a single, truly optimized crosshair sweet spot works for surround sound, just as it has worked for decades in the stereo two-channel context. This is not to say that if you are not precisely positioned in the sweet spot that the experience won't be satisfying. It just won't be "the best that it can be." Generally, a "sweet spot" can be shared by two people sitting next to each other on a love seat sofa without much compromise. But it won't support optimized holosonic imaging if you are seated anywhere other than that "sweet spot" position. This is also a truism for stereo-only listening.

As our readers are aware, the end result that the Home THX program sought to achieve was the sonic signature of a dubbing stage, the production environment in which motion picture soundtracks are mixed in. This environment has evolved over a hundred years as a mirror for a large audience motion picture experience. Movies have thus been produced to play in auditorium environments with audiences ranging from 200 to over 1,500 people. With the introduction of "Dolby Stereo In Selected Theatres" sound envelopment extended into the theatre as a mono signal, which produced the "fuzzy-wuzzy" characteristic. Even with split surround 70mm the end result became left wall and right wall "fuzzy-wuzzyness." Then came the Surround EX and the back wall contributed even more "fuzzy-wuzzyness." While sound designers and re-recording engineers may desire to create soundtracks with imaging-specific holosonic characteristics, the loudspeaker delivery systems in theatres are not capable of recreating such an experience for a commercial theatre's audience. Thus, the idea of "imaging" soundscapes has not been in the vocabulary of most filmmakers and exhibitors, nor for that matter the THX organization. This has been my criticism of the Home THX program since its inception.

The point I want to drive home is that we enthusiasts need to demand that the creative community and the deliverers of hardware and loudspeakers focus on creating the experience of theatre nouveau in the home, which is created on the basis of imaging-specific three-dimensional holosonic soundfields with lateral and depth perception all around our heads (and above for a spherical surround experience). This should be the goal, which not only will optimize movie soundtrack experiences but that of surround music, gaming, and HDTV as well.

A New Surround Sound Standard

Dear Gary:

I've never written you before but I was compelled after reading the June Issue 85.

Bravo!

We're using terms like 3-D and 360. We're talking about the varieties of multi-channel (home theatre, surround music, gaming, and HDTV), but while we've acquired the discrete channels in both hardware and software, we're not maximizing their potential. It's still really a Pro Logic® world.

I really like your honesty (for lack of a better term) on the "sweet spot." You're right, there's only one small area, and I'm sorry there's not room for everybody.

David Bales, HECA Marketing Manager-Audio, Pioneer Electronics,
david.bales@pioneer-usa.com

Editor-In-Chief Gary Reber Comments: Thank you for using common sense and recognizing the truth about the "sweet spot." But a great surround experience is still to be had in the close vicinity of the sweet spot and that is why imaging-specific-created soundfields will be applauded when experienced in the home in a system optimized to deliver such experiences.

Standard Issues

Dear Gary:

I think we have to pay attention to standards issues before things really start getting out of hand!

I started mixing music in surround with Stevie Wonder in the 1970s.

When 5.1 on disc was introduced in the 1990s, my partner, Brant Biles, and I were among the first to mix music in that format. We have since moved on to prepare some of the most prominent feature films for DVD at our company Mi Casa Multimedia. This includes the film soundtracks themselves as well as every sort of additional material. (For reviews and a list of titles, see www.micasamm.com.)

It has come to our attention that a recent RIAA/IFPI paper specifies a "recording industry position" for next-generation discs. This paper does not speak for us. When and where were the meetings held whence the conclusions of this paper were drawn? Who attended?

Specifically:

Why only six channels? This does not allow for a discrete center back surround (as already provided on DVD with DTS-DTS-ES), and there should be another channel provided in reserve, since it is likely that it will be needed in the future, perhaps as an overhead height channel. We have no argument with the sample rates proposed.

Why no mention of WM9 and DTS®? Was the paper written by Dolby employees? It



does not list a codec from any other licensor. We are in no camp and use every system. We have encoded thousands of hours in both Dolby Digital (AC-3®) and DTS (Coherent Acoustics®) and the superior quality of DTS is quite apparent to us. DTS has also announced lossless capability.

The RIAA paper does not adequately address the issues. We must get beyond favoritism and do what is best for the business and the consumer, and work toward inclusive and adaptable standards.

We have finally reached a stage in audio where the medium itself potentially outstrips the requirements of the content. Why limit what we are doing for the future?

Robert Margouleff, CEO, Mi Casa Multimedia, Inc., Los Angeles, California, robert@micasamm.com

Editor-In-Chief Gary Reber Comments: It always amazes me when "industry standards" are issued and there is no attributing to the authors of the standards, the actual individuals who position and advocate for the standard and write the issuing paper(s). Not only is this the case with the RIAA (Recording Industry Association of America) but it is also true of the surround music loudspeaker setup recommendations issued by the European standards organization ITU (International Telecommunications Union), another one of those mysterious organizations that don't involve cutting-edge practitioners in the creative community in the rendering of standards and recommendations.

In the early 1990s, the ITU began conducting research to determine optimum loudspeaker placement in a 5.1 configuration. This culminated in a document published in 1994 entitled "Recommendations For Multichannel Stereophonic Sound System With And Without Accompanying Picture" (Rec. ITU-R BS.775-1), which was largely accepted as a de facto industry standard. However, what is not common knowledge is that the ITU research, originally developed for television broadcast, was done almost entirely with classical music as source material, and that the back loudspeakers were characterized as "ambience" or "effects" loudspeakers only, and its recommendations were created well before the development of modern surround sound mixing methods, which give equal importance to all five main loudspeakers.

Now the RIAA and IFPI (International Federation of the Phonographic Industry), which is affiliated with the RIAA, the organization responsible for the world's largest music market, have issued a position paper entitled: "Recording Industry Position On

Technical Requirements For Next Generation High Capacity Optical Discs."

All this continues to create confusion for all concerned—from the creative community to the consumer.

I was able to obtain a copy of this "standards" paper and agree with your assessment of the lack of forward-thinking about what is needed to create imaging-specific holosonic spherical surround soundfields using more than 5.1 discrete channels.

With the pending launch as early as late 2005 of the Blu-ray Disc and/or HD DVD high-definition blue-laser optical disc formats there will ample provision at least eight discrete channels using all the accepted digital audio formats: Linear PCM (including Meridian Lossless Packing or MLP), DTS (lossy and lossless), Dolby Digital, and Windows Media 9 Professional (WMA Pro), a scalable lossy codec, which allows the streaming of up to eight channels of 24-bit/96 kHz audio data using variable bit rates ranging from 128 to 768 kilobits per second. Another codec is Windows Media Audio 9 Lossless. Such a single disc platform will have the capability to deliver the highest quality presentation of motion pictures, surround music, gaming, and HDTV programming.

You and Brant have been leading-edge practitioners of the state-of-the-art in music and movie surround sound mixing using your creative talents and engineering skills to create some of the very finest holosonic soundfields I have ever experienced. I will mention your outstanding work on the soundtrack repurposing and mastering for the DVD release of The Lord Of The Rings Trilogy. And, of course, I am looking forward to both Brant and you sharing with enthusiasts your approach to sound design for both music and, in particular, movie soundtracks on the upcoming 2004 Home Theater Cruise", October 24-31, 2004 (see www.hometheatercruise.com).

What we need is a convergence of the art forms to create a single imaging-specific three-dimensional soundfield standard setup for mixing and playback of movie soundtracks, surround music, gaming, and HDTV programming.

As you know, such an imaging-specific approach was discussed last year by the Producers & Engineering Wing of NARAS (National Association Of Recording Arts And Sciences www.grammy.com/pe_wing/guidelines/index.aspx) and on the 2003 Home Theater Cruise (Phil Ramone, Al Schmitt, Elliot Scheiner, Alan Parsons, Bobby Owsinski and yourself), as well as in numerous issues of Widescreen Review. Discussion of this topic is once again scheduled for the 2004 Home Theater

Cruise with the addition of Frank Filippetti, Michael Bishop, Eric Lindemann, Mike Sokol, Gary Marz, and Rick Dean of THX, Ltd. And, of course, you and Brant both deserve credit for having already mixed many music and movie projects in 6.1 using full-range direct-radiator loudspeakers (not THX-recommended dipoles) positioned to optimize holosonic imaging.

I see no problem with supporting at least eight discrete channels with 24-bit/96 kHz resolution, whether uncompressed, lossless, or lossy format.

I can understand the RIAA/IFPI leaving out WM9 because it is so new a codec but to leave out DTS is shameful! DTS has consistently proved excellent sonic legs with 48 kHz resolution and promises even superior performance with 24-bit/96 kHz and lossless resolution.

As well WM9 promises superb performance with 24-bit/96 kHz and lossless resolution.

With the capability of the new formats on the horizon, there is no excuse for limiting the choices for high-performance discrete multichannel audio. ■



More power to you

"...a completely new level of listening pleasure"

"...simply the best image quality I have experienced"

"...trust me on this: you don't know what you're missing"

-Doug Blackburn, WSR

Equi=Tech

"The Pioneer of Balanced Power"

Toll Free: (877) EQUITECH

www.equitech.com