Audio Engineering Society

Convention Paper

Presented at the 117th Convention
2004 October 28–31 San Francisco, CA, USA

Spatial Definition and the PanAmbiophone microphone array for 2D surround & 3D fully periphonic recording

Robert E. (Robin) Miller III[1] ©2004

[1] FilmakerTechnology, Bethlehem PA 18018 USA / www.filmaker.com [1]

ABSTRACT

Higher sampling rates are necessary for high spectral resolution, but it is higher angular resolution and precision that preserves source directionality, and therefore higher tonal/timbral quality of that source, termed *spatial definition*. In acoustic spaces that are extensions of musical instruments, voices, and sources of sound effects (for movies, virtual reality, training simulation), tonality is a major contributor to lifelike perception – but in audio reproduction, lifelike tonality is limited by the recording system. A surround microphone has been developed both for more precise 2D surround ("PanAmbio"), compatible with ITU 5.1 and stereo, and for "PerAmbio" 3D (with height) for the ultimate in tonal reality distributable using ordinary 6-channel media for either decoderless 2D replay or 3D with decoder and five additional speakers.

1. PURPOSE OF WORK; FORM OF PAPER

The goal of Ambiophonics [1,2,3,4] is more lifelike reproduction of sound. This ongoing pursuit was spawned by acknowledged issues of reproduction accuracy with stereo and stereo-based surround systems, including ITU 5.1 [5,6]. While 5.1 has improved over stereo as stereo improved over monaural, the situation is far from perfect for critical music or movie listening, or for realistic virtual reality (VR) and training simulation. Nevertheless, it is a "stereo world," and 5.1 is well accepted, so compatibility is of importance to users, as is extensibility as their needs change in the future. But beyond higher sampling rate or bit depth, more lifelike sound is the product of other perceptual qualities termed *spatial definition,* the first subject of this paper.

Second, for developing a microphone for *high spatial definition* recording systems, the objectives were:

- Lifelike spatial and enveloping sound;
- Accurate source localization and timbre;
- Relative ease of use with consistent results and future value of recordings for music, movie, broadcast, VR, & training simulator industries.

1.1. Lifelike, spatial, enveloping sound

We engage in discussion about "high resolution" audio, meaning using higher sampling rates, pushing upward the Nyquist limit of content frequencies and temporal resolution. But it may be argued that just as important is more precise spatiality, and resulting lifelike envelopment and correct tonality (timbral quality), of sources in acoustic spaces – termed "high spatial resolution" or "high spatial definition."

Contrasted with 2 channel stereo, 5.1 surround developed for home movie theaters offers advantages also for music reproduction. However, phantom images between 5 or more main speakers are highly imprecise, especially to the sides but well outside the *cone of confusion*. Add that arrivals from above and below are projected onto the horizontal plane of 5.1 speakers, placing listeners in the center of a circle, not the sphere of lifelike hearing. High spectral resolution alone cannot compensate if absent is high spatial resolution of indirect energy that may equal or exceed that of direct sound in ambient recordings.

Listening taste for recorded music and sounds seems to fall in one of two camps: *they are here,* or *you are there*. In the first case, sources are usually recorded using closely-placed microphones – one per instrument or section – and "dry" (inaudible venue acoustics). Often with this practice, the listening room acoustic is more reverberant than the recording, so any recording space "disappears" and the instrument "appears" inside a speaker box. Even if "artificial" reverberation is added, it is often sufficiently "disembodied" from the source so that the source still seems to be in the listening room. Although unnatural, such artificial "intimacy" is the learned hallmark of much popular music that has become de facto the audience's/market's acquired taste.

The second case – *you are there* – is the more natural state where the reproduction is, to the greatest extent possible, like the experience of being present at the recording. Now with the original venue's acoustics unmasked by the listening acoustics, the listener feels "transported" to the concert hall, movie location, or jet fighter cockpit. Rather than every recording sounding like the listening room (boring?), each recording sounds more like its real venue (you get to travel).

Assuming constant listening acoustics, the recording engineer has control over whether the *musicians are here* or the *listener is there* more-or-less by varying the venue-spatiality of the recording [7,8]. If the spaciousness of the recording is significantly less than that of the listening room, *they are here*; if the spaciousness of the recording is the greater, *you are there*. But this choice will not be best accomplished by superimposing disembodied artificial reverb – the recorded spaciousness must be perfectly matched reverberation (including early reflections), possibly by convolution with actual hall impulse responses, or directly recorded – the work of a spatial microphone technique that is the latter subject of this paper.

1.2. Accurate localization & timbre

Localization is important for more than direct sound. True, in life it is important to know when a bus is bearing down on you and from what direction. And matching screen direction is important for movies. But in acoustic spaces, preserving directionality of each echo arrival also preserves the timbre of that arrival. This is because of our unique spectral "coding" – or coloring – imprinted upon each direction by our individual HRTF. Including pinna effects, height as well as horizontal directions are interpreted in our brain by learned association between HRTF-filtered sound and experienced source direction.

Although our acuity for height is less (±10°) than in the horizontal plane (±1°) [9,10], when height is included in the timbral mix in our brain, we are in the center of not just a circle, but a sphere. Therefore the definition of envelopment and *spatial definition* is not just circular, but spherical. Each room reflection, arriving at ever-later times and colored ever-differently by our pinna to represent different directions, develops a composite timbre over the reverberation time of the room. When the source ceases, the sound collapses timbrally in the same complex order of tonal changes, as each reflection ceases over time and direction. Thus, musicians in the same room artfully form each note, phrase, even pause of their performance. Listeners in the same space regard subsequent reproduction as "correct" if this complex tonality-in-time is preserved, which means it must be preserved in directionality. Any height cues contributed by the listening room are invalid. Therefore, accurate recorded localization is key to the lifelike quality of sounds – and the "musicality" of music – and therefore key to high *spatial definition*.

These spatial qualities have been recognized in other advanced audio reproduction systems, such as (3D) Higher Order Ambisonics (HOA) [11,12,13,14] and (2D) Wavefield Synthesis (WFS) [15,16]. However elegant and promising, their requirements in terms of processing and number of channels may be impractical for home use for some time [17,18]. (On the other hand, "PerAmbio 3D/2D," discussed in 2.1 below, exhibits high spatial definition and is practical today, requiring just 6 channels distributable on available media (DVD-A, SACD, DTS-ES CD) and 10 speakers for fully periphonic 3D [19,20,21]).

Fig.1 - The PanAmbiophone recording the Greenwich Village Symphony, New York City. The experiment used no outrigger or spot microphones for ease of use capturing natural surround.

1.3. Ease of use – production & post

Errors in localizing phantom images were known to stereo's inventor, Alan Blumlein, but were necessary to the very way stereo worked (in transforming level difference into phase difference). The work-arounds are well-known to recording engineers, and are applicable in 5.1 as well. Capturing more accurate spatial signals in the original recording is inherent in the new microphone that is latter subject of this paper.

With stereo's acceptance, recording engineers developed ways to overcome problems of "hole-in-the-middle" ("bunching" at the speakers) and comb-filtering of important central (solo) sounds, developing use of such tools as equalization and spot microphones. As mentioned above, use of spot microphones has led to the popular taste for *they are here*, and has become standard practice for music recording and sound reinforcement, despite the other complexities of mixing, especially when done live. However, mixing any two or more microphones, each contributing different replicas in time of the same source, results in complex errors of tonality and imaging (due to comb filtering and time smearing). Advanced multi-microphone mixing techniques such as Room Related Balancing [22] can help mitigate these errors. But being able to record without (or with fewer) spot microphones would greatly simplify production or post-production and avoid degradation, and is inherent in the approach of the *high spatial definition* main-microphone described below.

1.4. Future value – compatibility with 5.1

As of this writing, it is still a "stereo world," although 2D surround reproduction such as ITU 5.1 is becoming more accepted, popularized by movies and DVDs. And surround music is a natural potential market for home theaters. Great future value awaits content producers, musicians, and home theater owners who demand a system that could produce stereo, 2D surround, and future 3D surround (with height) – and that exhibits backward/forward compatibility. The family of Ambiophonics adopts compatibility to the greatest extent in the pursuit of lifelike 2D and 3D reproduction, as described below.

1.5. Non-audio applications

Music and movies for home theater would benefit from the most lifelike reproduction. However, virtual reality (VR) for gaming and simulation for training are other important applications, especially if the system has sufficient accuracy of localization in 3-space (3D, with height), such as PerAmbio 3D [19,20,21, Appendix].

2. THE "FAMILY" OF AMBIOPHONICS

Ambiophonics embodies both prior and new art in the pursuit of lifelike audio. Briefly, the three "flavors" are:

- Ambiophonics ("Ambio") – 2-channel recordings, including many originally intended for stereo or binaural, with (optional) hall convolution [23]) – illustrated in Appendix;

- Panor-Ambiophonics ("PanAmbio" [24]) – 2D surround from 4-channel recordings (ITU 5.1-compatible with silent center) – illustrated in Fig.2 and Appendix Fig.B;

- Periphonic Ambiophonics ("PerAmbio" 3D (with height) full sphere surround (5.1/6.1-compatible from 6-channel encoded recordings Pat. pending) – illustrated in Appendix Fig.C, [19,20,21].

All three can be augmented with hall impulse response (IR) convolution [25,26], although both PanAmbio 2D and PerAmbio 3D record surround directly, correctly positioning hall ambience and sources at side, in back, and up/down in 3D. Information on all Ambiophonic approaches may be found in the Appendix and at www.ambiophonics.org and www.filmaker.com.

Spatial resolution/definition is addressed in all Ambiophonic approaches to stereo, 2D surround ("PanAmbio"), and 3D ("PerAmbio"). All preserve HRTF-related information captured in the recording and reproduce it using crosstalk-cancelled speaker pairs (stereo dipoles) that are closely spaced and in the general direction least confusing to the pinna. Over speakers, Ambiophonic reproduction works like "virtual headphones," reproducing binaural replicas of the original venue sounds, but more comfortably and with no inside-the-head sounds. In ITU 5.1, directionality (therefore tonality) is distorted, whereas in PanAmbio 4.0 it is precise within ±5° using double stereo dipoles front and back. A PanAmbio recording is made with a special microphone array (described below), has a lifelike quality, plays compatibly on 5.1/6.1 speaker layouts, and may if desired incorporate spot microphones and center channel processing.

Ultimately, reproduction with height is required to approach live 3D hearing, where arrivals from above and below also contribute to tonality due to individual pinna filtering. As said, in acoustic spaces, performers (musicians, vocalists, Foley artists, etc.) "play" the entire sphere of periphonic sound as extensions of their instruments, for which the integration of many head-timed and pinna-filtered arrivals define tonality also for the listener in the same space. Having learned this complex interplay from experience, individual listeners now expect the same interplay as their personal standard for live hearing – and recognize, in the absence of lifelike *spatial definition*, that it is "only a recording." So the ultimate step in realistic audio is full-sphere 3D.

2.1. *PerAmbio 3D/2D* – decoderless 5.1 or 3D

The ultimate in spatial definition is 3D (with height) reproduction, where the listener is at not just the center of a circle as with 2D surround e.g. 5.1, but at the center of the sphere of human hearing. "PerAmbio 3D/2D" (Pat. pending) [19,20,21] combines Ambiophonics and a modified Ambisonic soundfield, for the ability of each system in reproducing front stage and ambience, respectively, and delivers it on common 6-channel media (DVD-A, SACD, DTS-ES Discrete CD) to be replayed either on ITU 5.1/6.1 layouts without any decoder, or (from the same disc) in full 3D by adding a decoder and four or five or more speakers placed in positions programmed in the decoder. A "mode" for decoderless 5.1 2D mapping is chosen by the recording engineer, possibly changed in post-production, and selected upon replay by the user (automatically in metadata?). Alternatively or in combination to enhance ambience, hall sound may be convolved from hall impulse responses, either by producer or user.

The PerAmbio 3D/2D speaker layout (see Appendix Fig.C) is fully backward-compatible with 5.1 recordings by the listener simply repositioning back 26% of the speaker diameter, where speaker angles match the ITU standard. Height and surround speakers may be flexibly positioned and their coordinates programmed in the decoder. For other than the most critical music listening, or for movie-watching in 3D, up to six listeners can be accommodated in PerAmbio 3D/2D's broad listening area.

Recorded using a hybrid microphone based on the PanAmbiophone, below, PerAmbio 3D/2D is a multi-format-compatible system that allows performances to be preserved, and both producers' and users' libraries to have greater future value in a future 3D audio world.

3. *PANAMBIOPHONE* – A MAIN MICROPHONE

A "main microphone" approach [27,28,29,22], the "PanAmbiophone" – the second part of this paper – is integral to the concept of *high spatial definition*. Its purpose was to:

- Deliver 2D surround by direct recording ("PanAmbio") and exhibit uncompromised accuracy of localization and tonality;
- Offer compatibility with stereo and ITU 5.1 for cinema, auto sound, broadcast, home theater;
- Simplify surround production/post-production for music, movies, VR, & training simulation using a main microphone approach that obviates the need for many spot microphones;
- Form the basis for a 3D (with height) microphone and a multi-format-compatible system for lifelike reproduction (PerAmbio 3D).

The original PanAmbiophone design [24] married two sphere microphones, for HRTF qualities, front and back of an acoustic barrier – see Fig.4 & 5. The barrier was a sandwich of gypsum insulator between two absorbers each 2 inches (50mm) thick – and was big and heavy. Two persons were required to mount and raise the array on a substantial stand. Nevertheless, the results, using Schoeps microphone elements, were encouraging, as indicated by the subjective comments and performance curves in [24] and below. This experience evolved a new design intended for the more useful recording in acoustic spaces.

The newest PanAmbiophone, shown in Fig.9, weighs 10lb (5kg), and may be mounted on a stand or suspended. Approaching the "perfect omni," the array may be positioned farther from sound sources than is conventional, beyond the critical radius, with results that seem more "present" than expected, obviating use of many spot microphones. Played on ordinary 5.1 home theater speakers, the results are ambient and tonally natural. Played using dual crosstalk-cancelled speaker pairs – PanAmbio as in Fig.2 – the measured results are "correct" localization in the horizontal plane. Played using PerAmbio 3D/2D speakers (see Appendix Fig.C), the results are subjectively highly lifelike, as described by auditioners at the AES 24th International Conference in Banff, Canada, June 2003 [21,30].

3.1. Directional head-shaped microphone array

The PanAmbiophone has evolved in prototype form over several generations, beginning with two sphere microphones [31] separated by an acoustical barrier, as shown in Fig.4 & 5. The original objectives [24] were:

- Easy to use, HRTF-related main microphone
- "Perfect omni" resp. 5Hz~30kHz around 360°
- Directly records 360° surround (horizontal 2D)
- Independently controllable front & back stages
- Stable images around 360°; accurate localization
- Lifelike tonality
- Compatible with ITU 5.1 (center silent)
- Basis for direct recording of 3D "full sphere"
- Optional 3D ambience by convolution

Subjective testing [24] contrasted recordings played both on ITU 5.1 layouts and over dual stereo dipoles (PanAmbio 4.0) using crosstalk cancellation by DSP, illustrated in Fig.2. In PanAmbio, antiphonal sources, audience sounds, and reflections were heard imaging within ±5° except in the human "cone of confusion."

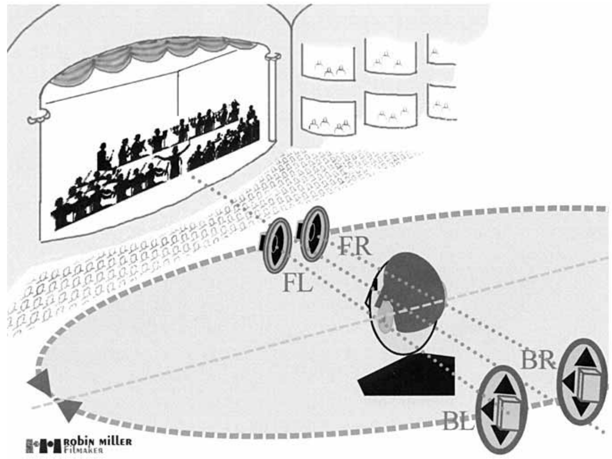

Fig.2 illustrates Panor-Ambiophonic ("PanAmbio") layout using two closely-spaced speaker pairs for 360° horizontal (2D) surround reproduction for home theater music and movies.

The original design simulated four quasi-cardioid microphones, limited by the cutoff frequency of an acoustical barrier, as in Fig.3. Unlike real cardioid microphones, or the approach by Bruck combining omni and bi-directional elements, the frequency response using four omnis was flat to below 20Hz, along with excellent high frequency polar response around 360°.

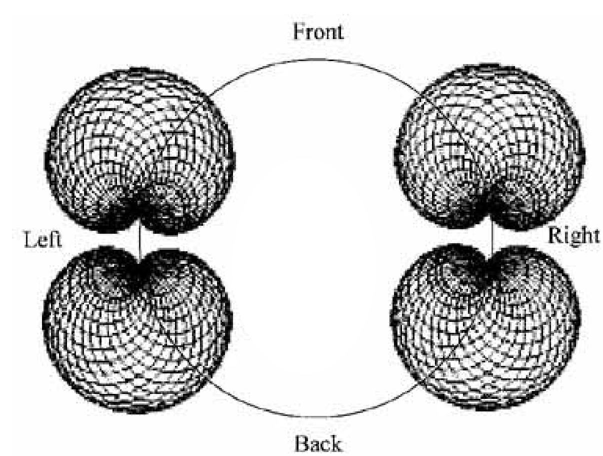

Fig.3 – Model of the original PanAmbiophone – four ideal cardioids tangent to a head-shaped baffle.

Each pair of microphones was head-spaced at the surface of an acoustically rigid sphere. Simultaneous recordings were made for comparison using OCT [22] for 5.1 reproduction and Ambisonics for a total of 12 microphones, shown (with barrier removed) in Fig.5.

Fig.4 – The first PanAmbiophone with OCT for comparison experiment. Barrier hides second sphere for rear stage.

Fig.5 - Original PanAmbiophone using two spheres (barrier removed), OCT, and Soundfield microphones for comparison.

The directionality of each pair, front and back, measured with band-limited pink noise, demonstrated side-to-side separation typical of the sphere microphone – approx. 10dB. The baffle also created a front pair to back pair separation of approx. 10dB. However, compared to real directional (pressure gradient) microphones, the response was again flat to lowest frequencies and remained directional to the limit of the size of the baffle. The baffle was scaled consistent with HRTF cutoff of approx. 700Hz.

Directional measurements are plotted in Fig.6 & Fig.7. With axes of the microphone elements at ±90° directly left and right, the high frequency polar response over all 360° was more truly omni-directional, making more distant positioning possible. Still more "perfect omni" polar high frequency response would be realized in a new, directional PanAmbiophone, described in section 4.0 below.

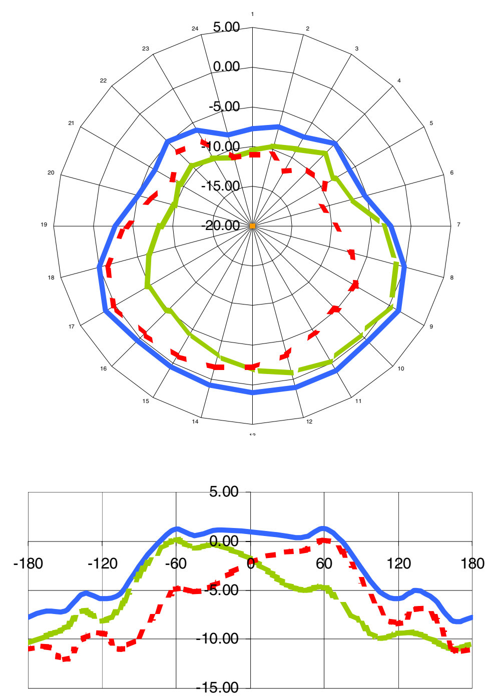

Fig.6 - Plots side-to-side response to band-limited pink noise for one microphone pair (0° is down on polar graph). Whether used front- or back-facing, the summed response (solid line) shows a polar characteristic that is flat ±1dB across each stage.

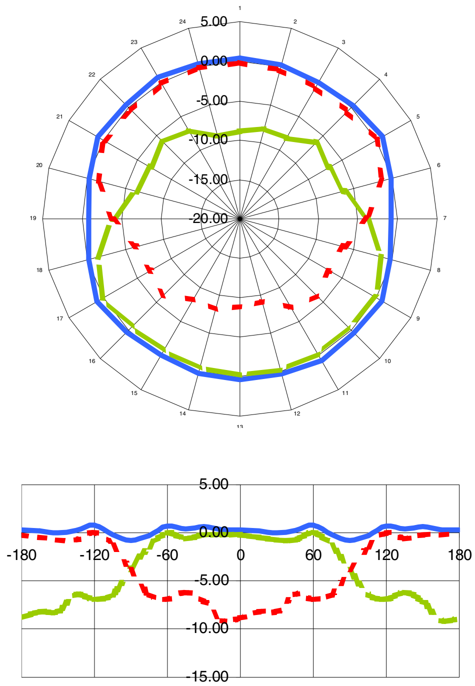

Fig.7 plots front-to-back response to band-limited pink noise (0° front is down on polar graph). Summed surround response 360° (solid line) shows a polar characteristic that is flat ±1dB.

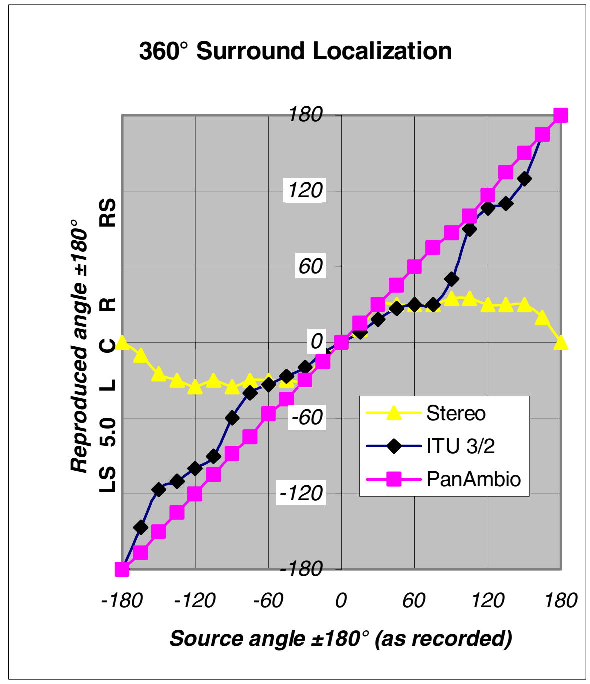

Fig.8 – Plot of informal tests for localization of speech signals at 15° intervals around 360° comparing PanAmbio, 5.1, & stereo. The PanAmbiophone replay consistently localized within ±5°.

The PanAmbiophone's benefits also are directly applicable for 2-channel stereo or Ambiophonic 2.0, whether mixed live or in post-production, or down-mixed in the user's player or receiver. The array is positioned by the recording engineer for optimum recording angle and imaging the front stage. The rear staged is then mixed in for balancing stage sound with ambience, typically at a level lower by 1½ to 3 dB relative to multi-channel use.

In a comparison test with 5.1 and stereo, in Fig.8, speech signals were recorded at 15° intervals around 360°. Except near the human cone of confusion, localization results for PanAmbio were consistent ±5°. This accuracy implies: 1) preservation of HRTF-based cues; 2) front-back discrimination that is controllable and allows more distant (natural) pickup; and 3) preservation of arrival direction necessary for correct timbre for each listener, using their own HRTF including pinna. The result using the PanAmbiophone, therefore, is not just raw, artificial spatialization, but naturally *high spatial definition* and lifelike reproduction.

4. DIRECTIONAL PANAMBIOPHONE

The results in Fig.6 through Fig.8 showing 360° polar characteristics and localization approached what was thought to be the ideal for horizontal (2D) surround. In practice, this characteristic is applicable mostly to anechoic environments, such as outdoors. Subjectively, good results were obtained recording a passing parade and natural ambience, such as for movies and nature recordings. However, in acoustic space, the original design reproduced too much ambience in front, as is common with sphere microphone recordings (where the balance of stage-to-hall is fixed by positioning the microphone, and is not electronically controllable, live or in post-production, as it is with the PanAmbiophone).

Especially for staged presentations, a different characteristic may be more desirable – one that focuses the front microphone pair more on the stage, allowing a deeper range of placement of the array farther from the source, such as an orchestra, than is possible with the original PanAmbiophone. Therefore, a new design was conceived that reproduced the 120° stage in front and the 240° hall in back, while still being able to control stage-to-hall balance, live or in post-production.

4.1. Focused surround microphone for interior recording

Design criteria for the latest in the PanAmbiophone evolution intended for use in recording staged performances in acoustic spaces were:

- Ellipsoidal approximation of human HRTF [32]
- Simple barrier directionality (minimal artifacts)
- Front stage 120°, independent back stage 240°
- Control of ceiling reflections in front stage
- HRTF-related ITD, ILD shadowing (no pinna)
- Minimal degradation of mic. Impulse Response
- Frequency response compensatible to 30kHz
- Boundary layer operation (no comb-filtering)
- SNR +6dB over microphone elements
- 4.0 outputs (48V phantom power)

Achieving directional characteristics using mechanical boundaries has advantages in that small pressure (omni-directional) microphone elements are used, with resulting SNR equivalent to larger microphones. With small diaphragms, high frequency phase response more closely approaches the ideal, along with the good low frequency response of omni-directional elements. However, designing the complex baffle had risks, and had especially to ensure that no reflections would smear the high frequency impulse response and cause comb filtering. As in the original PanAmbiophone, the four microphone elements simulate human "ears," but are now positioned on a single head-shaped ellipsoid [32]. Now tangent to an acoustically hard surface, each diaphragm acts as a barrier microphone, with 6dB acoustic gain above cutoff frequency, resulting in a 6dB greater SNR. So as to avoid unwanted reflections causing comb filtering of direct sounds, each element had to be tangent to three planes – side baffle, top baffle, and the plane tangent to the ellipsoid. Unlike the thick barrier of the original PanAmbiophone above, in the new design, each pair of elements front-to-back had to be within 1 inch (25mm) to correlate in the listening acoustic "mix" as though a single "ear." The design is illustrated in Fig.10 and the prototype is shown in Fig.9, with an attached discrete soundfield array for PerAmbio 3D recording (with height).

Fig.9 - New design of the PanAmbiophone. Two pairs of small pressure microphone elements are coincident at "ear" positions. A discrete soundfield array is mounted atop – a total of eight recording channels for PerAmbio 3D (with height).

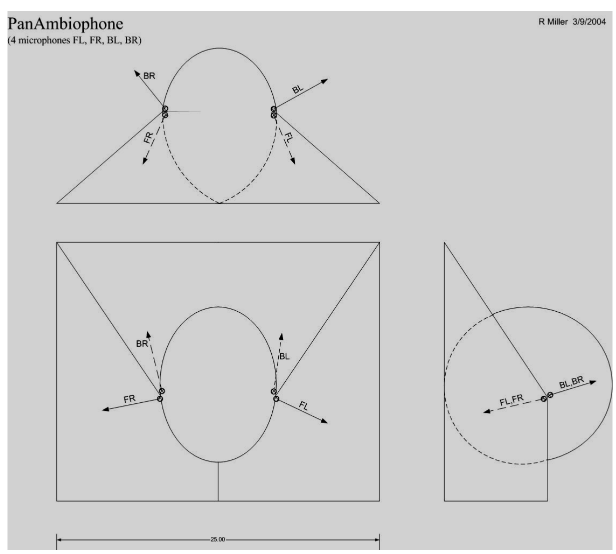

Fig.10 - Top, front, and side engineering views for the directional PanAmbiophone show ellipsoidal head-shape, four microphone elements, and typical source angles.

The directional PanAmbiophone and PerAmbio 3D/2D, both 5.1 compatible, are Pat.pending.

5. MEASURED PERFORMANCE

In Fig.11 through Fig.19 are plots of unequalized impulse and high frequency response 3kHz~30kHz of the prototype 4-channel PanAmbiophone, tested with a 1μs acoustic impulse (spark). The data provided clues to any errant reflections in the complex structure that might smear the response in time, and, therefore, degrade sonic performance. Of the total measurement series every 15°, the highlights are presented here.

In Fig.11, the impulse response (IR) of the front channels shows a risetime of 13μs with good settling. For frontal sources – usually the most important – this IR response is preferable to that typical of sphere microphones, where optimum performance occurs for sources at ±90° directly left or right [33-Fig2]. For reference, the high frequency response of the ¼ inch (6mm) microphone elements before mounting in the PanAmbiophone is in Fig.12, showing the response to be –2dB at 30kHz. Fig.13 through Fig.16 show, respectively, the unequalized high frequency response 3kHz~30kHz of front channels for a source at 0° directly front, 30° ipsolateral, 30° contralateral (showing lower level and greater tilt typical of the head shadow), and 60° ipsolateral. Fig.17 through Fig.19 show, respectively, the back channels for a source at 60°, 120°, and 180° directly back.

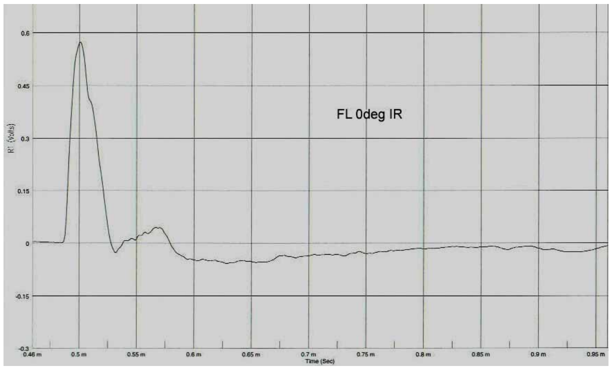

Fig.11 - Impulse response of PanAmbiophone front channels using 1μs spark. Measured risetime is 13μs with rapid settling. Horizontal divisions are 50μs.

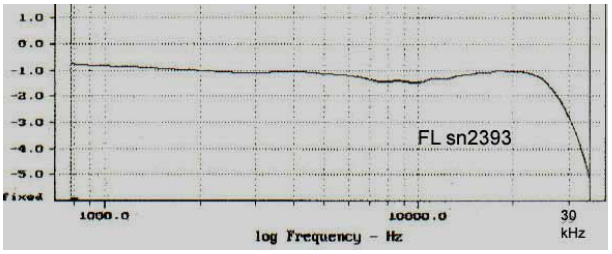

Fig.12 - Frequency response 3kHz~30kHz of the Earthworks elements before mounting in the PanAmbiophone shows response -2dB at 30kHz.

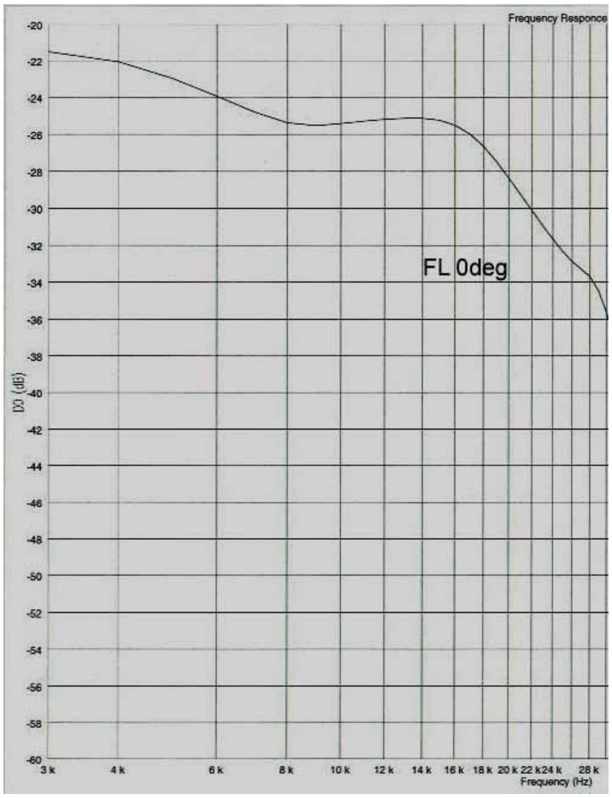

Fig.13 - Unequalized frequency response 3kHz~30kHz of the PanAmbiophone front channels for source directly front (0°). Horizontal divisions are 2kHz with 30kHz at the extreme right.

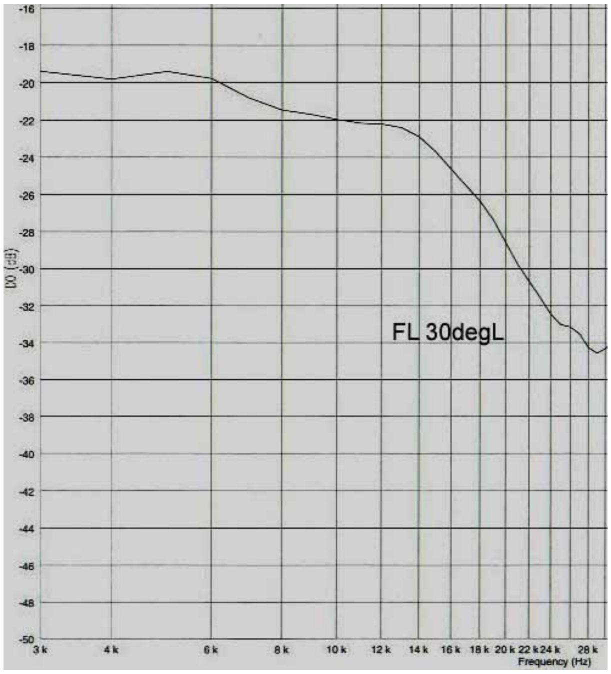

Fig.14 - Unequalized frequency response 3kHz~30kHz of front channels for a source positioned ±30° (ipsolateral).

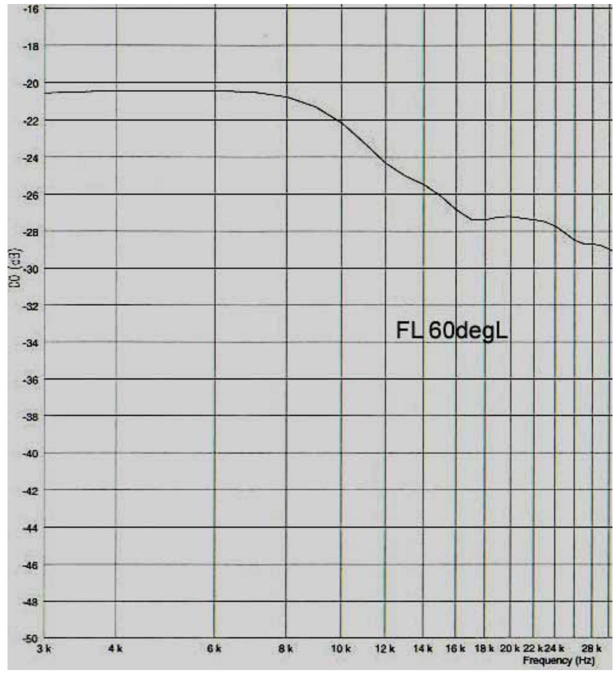

Fig.16 - Unequalized frequency response 3kHz~30kHz of the PanAmbiophone front channels for a source positioned ±60° (ipsolateral).

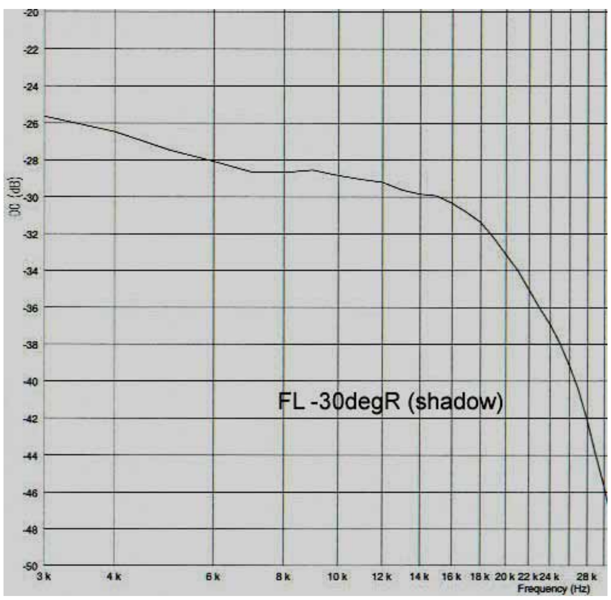

Fig.15 - Unequalized frequency response 3kHz~30kHz of the PanAmbiophone front channels for a contralateral source positioned ±30° - shows response in "head shadow" of the microphone on the opposite side from the source.

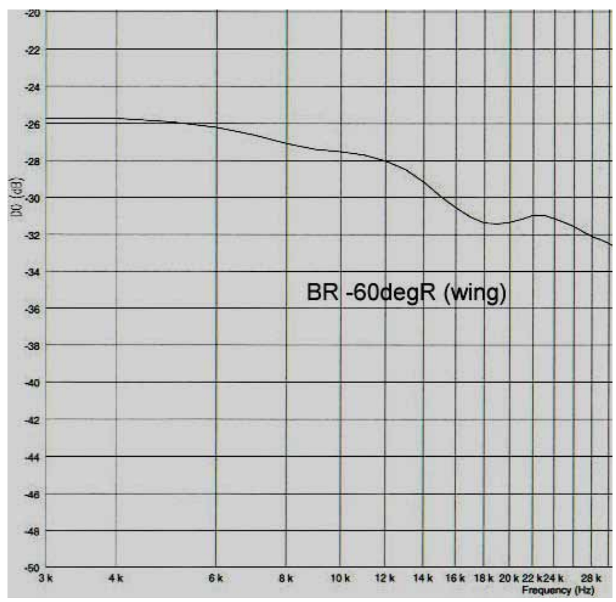

Fig.17 - Unequalized frequency response 3kHz~30kHz of the PanAmbiophone *back* channels for a source positioned ±60° (ipsolateral)..

Fig.18 - Unequalized frequency response 3kHz~30kHz of the PanAmbiophone *back* channels for a source positioned ±120° (ipsolateral).

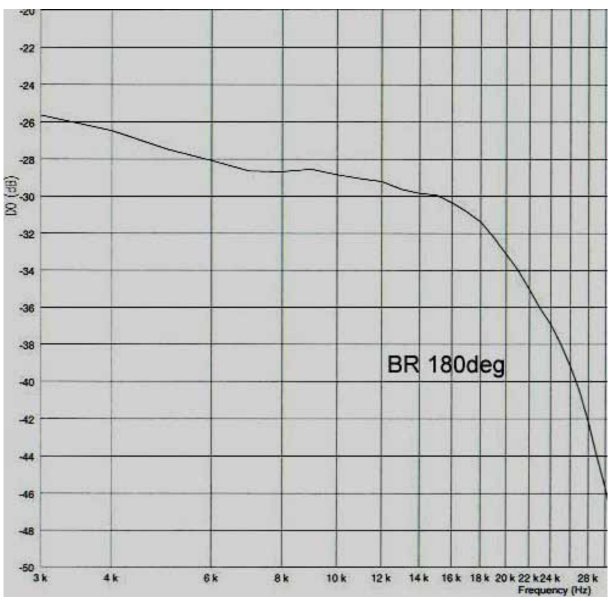

Fig.19 - Unequalized frequency response 3kHz~30kHz of the PanAmbiophone *back* channels for a source positioned 180° directly back.

In the total measurement series every 15°, only one significant reflection error was found – measuring 0.235µs at source angle 45° on the contralateral microphone, down -9dB. The problem has been addressed with dampening compound within the structure. Also, using equalization, the high frequency response of the array has been corrected to within 2dB at 16kHz around 360° for a "perfect omni" result overall. Perfectly omni-directional high frequency polar response implies positioning the microphone at distances from the source that approach natural hearing, yet with with "reach" and "presence" obviating spot microphones. Below the cutoff frequency of the baffle, all four microphone signals are sufficiently correlated so that their signals add to a flat low frequency response. The overall frequency response of the PanAmbiophone is 5Hz to 30kHz. It is consistent to the extent possible with the *spatial definition* concept defined earlier in this paper.

6. TESTED APPLICATIONS

Numerous recordings have been made with each PanAmbiophone during its evolution, ranging over many musical genres (e.g. opera,, orchestral, organ, big band, brass quintet, bluegrass in a club, etc.) and ambient sounds (e.g. playgrounds, parades, etc.). In many cases, additional microphones representing INA [29], OCT, and soundfield arrays have been recorded simultaneously for comparison, such as for the multi-format demonstration led by the author at the AES 24th International Surround Conference in Banff, Canada, June 2003 [30,21]. For all demonstrations and reported results, no equalization, level compression, or other effects were used. Results have been more or less gratifying, leading to further improvements of both the PanAmbiophone and the technique for using it.

The discovery of the need for different polar characteristics for outdoor and indoor use has been discussed. The latest PanAmbiophone, designed for staged performances inside acoustic spaces, has been used for recordings of an oboe trio in a studio, chamber orchestra (10 players) in an auditorium, choir in a church, and orchestra in a hall with increasing success. Each of these sessions is briefly described:

1) In a 400 seat auditorium with a 10 piece chamber orchestra, the PanAmbiophone was placed at about the critical radius of the room (approx. 5m) with players arranged within a relatively wide 100° recording angle. The

reproduction stage was equally wide using a PanAmbio speaker layout (Fig.2, Appendix). Clinging to many years of traditional stereo and surround recording experience, this early experiment erred in too close a placement of the new microphone and so, with front and back channels in calibrated balance, proved overly dry. However, in post, the front-back balance was "fixed" by independent control of the front and back microphone pairs. Since it favors frontal sound (much as a cardioid increases working distance over an omni), the new microphone should have been used somewhat beyond the critical radius. A byproduct is that the array can be suspended above audience (or camera) sightlines.

2) The oboe trio in a 500m^3 soundstage with acoustics set for "live" was recorded beyond the critical radius (at 3m cf. 2.5m) with the trio subtended within a 100° recording angle. Replay in PanAmbio preserved this angle (in stereo or 5.1 it was, of course, limited to 60°) and resulted in good balance as calibrated between direct and early reflections from all around the space. The recording exhibits clarity and intimacy and is popular with recording engineers at demonstrations. Upon hearing a replay in the control room but sensing that the sound replicated the soundstage, a musician involved exclaimed "Finally, it's my sound." Listeners who represent consumers, after hearing the piano played in the soundstage, then hearing the recording in the control room, report "It sounds like *that* room, not like *this* room." – which can be interpreted as meeting the objective of *high spatial definition*.

3) For the 2nd anniversary concert of 9/11 in a church at Ground Zero in New York City (see Fig.20), the Rutter Requiem was performed by the professional chorus Seraphim and eight instrumentalists within a recording angle of 90° of the PanAmbiophone placed beyond the critical distance at 6m. Microphone levels were calibrated using band-limited pink noise. (While artistically quite satisfying, high RF interference from countless security radios marred the recording – the problem has been rectified since.) The very moving music, accompanied by the tolling of a large bell outside for each person killed, was faithfully reproduced – the choir, ensemble, and the church's acoustics in precise balance and tonally natural.

4) Columbia University Chapel in New York City was the venue for the Greenwich Village Orchestra playing Beethoven's 9th Symphony (see Fig.1). 80 players, 60 choristers, and 4 soloists were recorded beyond the critical distance (4.5m from soloists) within a wide 120° recording angle. Played in PanAmbio, the soprano and bass-baritone are heard nearly that wide – about 110° apart. The chorus, positioned deep and underpowered cf. the orchestra to the live hearer, is reproduced exactly so. However, in the mix, both the ambience and resonance of the male voices was improved by adjusting the back-to-front balance. The orchestra and chorus and their conductors have expressed praise for the artistic integrity of the recording. (Recording engineers agree, but must hear beyond the not-quite-commercial quality of the semi-professional performance, recorded live before a student audience – with cell phones.)

Due to the 120° front stage and 240° back, the new PanAmbiophone trades absolute directionality for good balance and ease of use inside. Outside, as expected, 2D space is warped when played in PanAmbio, as the 120° front stage becomes 180° in front of the listener, and the 240° back stage becomes 180° in back of the listener. While not exhaustively measured as of this writing, suffice it to state subjectively that it works as designed: a passing parade, marching in a straight line, was warped into a noticeable V shape, starting from back-right and ending back-left. On a playground, kids playing in a circle are heard to speed up around back and slow down in front. The original PanAmbiophone would have preserved these angles within ±5° around 360°, as plotted above. The choice of pattern is up to the recording engineer.

7. SUBJECTIVE RESULTS, FUTURE WORK

Anecdotal comments by the musicians involved and by representatives of consumers at demonstrations are briefly mentioned in the above section. Some expressed themselves with eye-popping, jaw-dropping, and requests for owning their own Ambiophonic system. A small minority report they cannot hear the wide stage or localize sources within it – Ambiophonics researchers believe there is some learning required, similar to acquired conditioning to conventional stereo/5.1.

Engineer-auditioners at the AES 24th Surround Conference in Banff, Canada, June 2003 heard recordings made with the new PanAmbiophone and generally commented positively [21], especially about PerAmbio 3D over just 10 small speakers [19,20,21]. Some report a sensation of "pressure" near the Ambiophonic mid-plane, only on which crosstalk cancellation occurs. Hybrid PerAmbio 3D is more forgiving in this regard, presenting credible envelopment for a broad listening area that, in Filmaker Technology's listening room, accommodates six persons (Appendix Fig.C). Still, critical listeners would want to be at the Ambiophonic "focus" for the most accurate sound.

For future experimentation and demonstration, it is hoped that a record label will permit a simultaneous recording by the author using the PanAmbiophone for comparison of the producer's normal recording with PanAmbio 4.0 and PerAmbio 3D/2D. Both of these "flavors" are 5.1-compatible directly and can be released on any "shiny disc" media. PerAmbio 3D/2D requires 6-channel media e.g. DVD-A, SACD, or DTS-ES Discrete CD. Also planned are refinements to the PanAmbiophone structure and equalization for use in commercial recording, movies, or broadcast.

Fig.20 – PanAmbiophone at right recording 9/11 requiem near Ground Zero, New York, in 5.1-compatible PanAmbio surround.

8. CONCLUSIONS

More than just high definition in terms of higher sampling resolution, *spatial definition* calls attention of recording engineers to the need to address qualities that contribute to higher perceived reality: 1) Achieving more natural, lifelike audio reproduction both in 5.1 and other formats such as those of the Ambiophonic family. 2) Solving problems associated with phantom images, multiple uncorrelated spot microphones, and smeared temporal response of microphones and speakers. And 3) Raising the bar of audio reproduction to true 3D in order to approach live hearing. The ongoing pursuit of Ambiophonics is to achieve this higher *spatial definition* not just for critical music listeners, but for *all* listeners of music, movies, VR, and training simulation.

One flavor of Ambiophonics is 5.1- compatible *PanAmbio*phonic ("PanAmbio") 2D surround (see Fig.2, Appendix). High spatial resolution/definition PanAmbio recordings that are compatible with 5.1 are made with an HRTF-based four-channel microphone array termed a "PanAmbiophone" intended:

- For ITU 5.1 surround, stereo, or…
- …*PanAmbio*phonic (double stereo dipole)
- Accurate 360° locating
- Lifelike tonality
- 5~30kHz frequency response, flat ~360°
- 4.0 output (5.1/6.1 with processing)
- Extensible to 5.1-compatible *PerAmbio 3D*

In addition to contributing to *high spatial definition*, the PanAmbiophone features ease of use, obviates need for spot microphones, and allows positioning farther from the source (e.g. an orchestra) than other main microphone approaches – even well beyond the critical radius, thereby more closely emulating human hearing in ideal positions in live venues (and helping to compensate for its size in sightlines and camera angles).

Discussed are the PanAmbiophone's design evolution and measured performance from 5Hz to 30kHz, accurate localization around 360°, uncompromised impulse and phase response, choice of polar characteristics for outdoor or indoor use, and uncompromised SNR, along with compatibility considerations related to directionality for reproduction using either double stereo dipoles (PanAmbio), or ITU 5.1/6.1 speakers. Even placing the array farther from the source produces a wide (120°) front stage in PanAmbio with a natural balance of depth of front-and-center sources to more distant ones without spot microphones. When replayed on 5.1 layouts, side and rear directionality and therefore tonality and spaciousness compare favorably with ITU-intended recordings (not meaning re-panned "multi-mono"). The result in either system is a useful instrument for natural-sounding reproduction – *high spatial definition* to use along with high spectral resolution.

The directional PanAmbiophone is described as a tool for controlling indoor ceiling reflections and front-to-back imaging for Ambiophonic, 5.1/6.1 and stereo downmix. Generally requiring no spot microphones, the PanAmbiophone is easier to use, and positioning of the array with regard to both recording angle and critical radius and the capability of balancing front and back stages in post-production are discussed. Applications are explored in typical venues for recording and broadcast and for classical, jazz, pop, and movie genres for 2D surround, for home theater use. VR and training simulation are other important applications that benefit from the PanAmbiophone's localization accuracy.

For future value in multiple formats including 3D of sessions recorded today using the PanAmbiophone, PerAmbio 3D/2D uses a "mode" for decoderless 2D mapping chosen by the recording engineer, possibly changed in post-production, and selected upon replay by the user (automatically in metadata?). For monitoring and in home theaters, height and surround speakers may be flexibly positioned and their coordinates programmed in the decoder. Optionally, to enhance ambience, hall sound may be convolved from hall impulse responses, either by producer or user. Recorded using the PanAmbiophone in 5.1-compatible 2D or PerAmbio 3D, performances are preserved, and producer and user libraries will not become obsolete.

PanAmbio 2D surround, PerAmbio 3D/2D, and the PanAmbiophone have been realized, experimentally tested, and subjectively evaluated – ready for recording engineers for easy use, psycho-acoustic correctness, and investment in audio content intended to have greater future content prized for high spatial definition.

9. ACKNOWLEDGEMENTS

The author is grateful to Ralph Glasgal for his inspiration and wisdom and to the Ambiophonics Institute for support of the author's work. Additional support is by the University of Parma with special thanks to Angelo Farina, Enrico Armelloni, and Anders Torger; Alex Khenkin of Earthworks; Geoff Houser, Craig Greenwood and Brad Frikkers for fabricating microphone prototypes; and the many musicians who performed for the test recordings. All trademarks are those of their owners.

Appendix – The Family of *Ambiophonics*

The Ambiophonics family, championed by Ralph Glasgal [23], includes:

- Ambiophonic 2D – 2.0 channels to 2 or more speakers (+SW) with hall convolution (Fig.A);
- PanAmbio (panor-ambiophonic) 2D – 4.0/4.1 channels to 4 or more speakers (+SW) (Fig.B);
- PerAmbio (periphonic-ambiophonic) 3D/2D – 6.0/6.1 chan. to 10 or more speakers (+SW) (Fig.C)

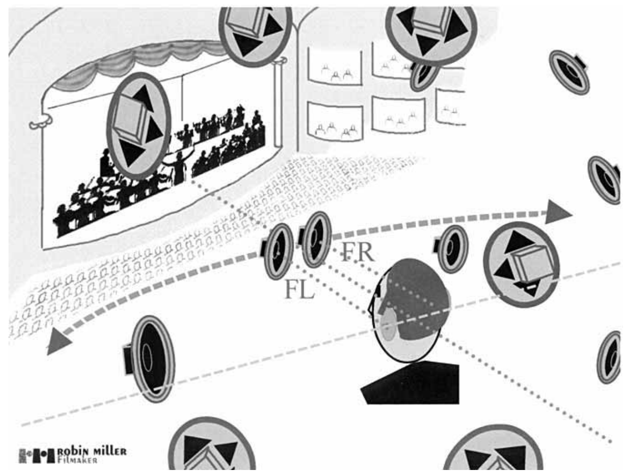

Fig.A - **Ambiophonic 2.0** turns stereo "inside-out," reproducing a 120° wide, natural sounding front stage using 10~20°-spaced speakers. Add 3D ambience speakers using impulse response convolution. Stereo recordings play compatibly in Ambio.

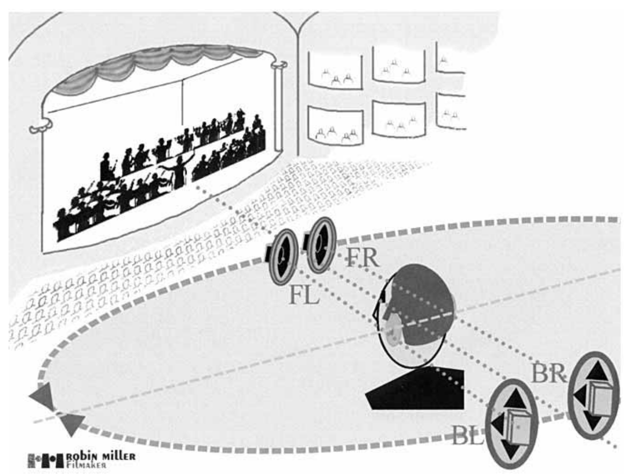

Fig.B - **PanAmbio 4.1** duplicates the wide front stage in back, "pulling" sides to the edges of the "cone of confusion" of human hearing. 5.1/6.1 music and movies play compatibly in PanAmbio; and 2D surround recordings made with the PanAmbiophone play compatibly on 5.1/6.1 systems.

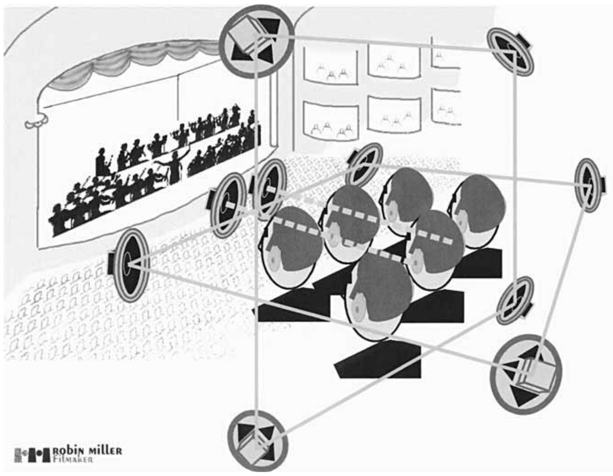

Fig.C – **PerAmbio 3D/2D** (Pat.pend.) plays both Ambiophonic and 3D (with height) recordings using 10 speakers. Height speakers may be flexibly positioned and their coordinates programmed in the decoder. Sitting back 26% of the speaker diameter results in standard ITU 5.1 angles for compatibility.

Ambiophonics works best with pinnaless binaural records, but enhances many stereo releases (even panned "multi-mono"), adding 3D ambience by hall convolution, and presenting a wide, accurate stage with no pinna confusion for central voices. Also, compatible with 5.1/6.1, specially recorded PanAmbio (2D) and PerAmbio 3D/2D releases (Pat.pending) allow music lovers and audiophiles a future path to greater listening precision in both localization and tonality.

Because they are more precise, all Ambiophonic playback requires good listening acoustics [34]:

- Acoustically treated room with RT less than recording venue and symmetrical layout of the speaker sphere, listeners seated near center;
- Two (Ambio) or four (PanAmbio) high quality speakers, plus 8+ satellite-grade speakers (total 10 speakers or more plus subwoofers – yes two! [35]);
- DVD-A/SACD/DVD-V/CD DTS-ES Discrete player with 6 full range channels;
- Decoder* that selects the transformation mode;
- Crosstalk-canceller* for front speaker pair;
- Bass manager accommodating 10+ speaker feeds (preferably two channel [35]);
- 2 higher and 8 or more lower power amplifiers;
- Calibration of channels at listening position within ½dB using an SPL meter, pink noise.

* possible on a DSP chip, now in prototype form [20].

Height and surround speakers may be flexibly positioned and their coordinates programmed in the decoder. A number of listeners not near any one speaker can enjoy enveloping 3D, but only one or two on the median plane will hear precise, 120° wide front localization. Mixes intended for ITU 5.1/6.1 can be enjoyed by simply moving back 26% of the speaker diameter [21].

REFERENCES

1. R. Glasgal, "Ambiophonics: Achieving Physiological Realism in Music Recording and Reproduction," Proceedings of AES 111th Convention, preprint 5426.
2. A Farina, R Glasgal, E Armelloni, A Torger, "Ambiophonic Principles for the Recording and Reproduction of Surround Sound for Music," Proceedings of AES 19th Int'l Conf., Schloss Elmau, Germany, 2001.
3. R. Miller, "Ambiophonic Demonstration: AES Bavaria 2001" 2-chan. CD with print insert, and streaming Ambio demonstration, "Surround" pages, www.filmaker.com.
4. R Glasgal, "Surround Ambiophonics Recording and Reproduction," Proceedings of AES 24th Int'l Conf, Banff, Alberta, Canada 2003
5. ITU-R BS.775.1: "Multichannel Stereophonic Sound System With and Without Accompanying Picture" (Geneva, 1992-4).
6. T. Holman, "5.1 Up and Running," Focal Press, ISBN 0-240-80383-3.
7. F. Rumsey, T McCormick, "Sound and Recording," Focal Press, ISBN 0-240-51680-X, pp 382-388
8. D. Griesinger, "The Psychoacoustics of Listening Area, Depth, and Envelopment in Surround Recordings and Their Relationship to Microphone Techniques," Proceedings of AES 19th Int'l Conf., Schloss Elmau,
9. Blauert, Jens, "Spatial Hearing: Psychophysics of Human Sound Localization," ISBN 0-262-02413-6, MIT Press
10. J Barbour, "Elevation Perception: Phantom Images in the Vertical Hemisphere," Proceedings of AES 24th Int'l Conf, Banff Canada 2003
11. M. Gerzon, "Ambisonics in Multichannel Broadcasting and Video," JAES Vol. 33 No.11, 1985, pp 859 – 871T.
12. J Daniel, R Nicol, S Moreau "Further Investigations of High Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging," AES 114th Int'l

Convention, Amsterdam, The Netherlands, 2003, preprint #5788.

13. J Daniel, S Moreau, "Further Study of Sound Field Coding with Higher Order Ambisonics," AES 116th Int'l Conv, Berlin, May 2004, preprint #6017
14. Malham, D.,http://www.ambisonic.net/
15. D de Vries, "Experience with a Wavefield Synthesis System Based on Multi-actuator Panel Loudspeakers," as presented at AES 24th Int'l Conf, Banff, Canada 2003
16. G Theile, H Wittek et al, "Potential Wavefield Synthesis Applications in the Multichannel Stereophonic World," Proceedings of AES 24th Int'l Conf, Banff, Canada 2003
17. P Craven, R Stuart, R Wilson et al, "Hierarchical Lossless Transmission of Surround Sound Using MLP," Proceedings of AES 24th Int'l Conf, Banff, Canada 2003.
18. ATSC Document A/52A "ATSC standard: Digital Audio Compression (AC-3), Rev. A" 8/21/2001, www.atsc.org/
19. R Miller, "Compatible PanAmbiophonic 4.1 and PerAmbiophonic 6.1 Surround Sound for Advanced Television – Beyond ITU 5.1," SMPTE 144th Technical Conf. & Exhibition, Pasadena CA USA, October 2002
20. R Miller, "Transforming Ambiophonic + Ambisonic 3D Surround Sound to & from ITU 5.1/6.1," Proceedings of AES 114th International Convention, Amsterdam, The Netherlands, March 2003, preprint #5799.
21. Robert E. (Robin) Miller III, "Scalable Tri-play Recording for Stereo, ITU 5.1/6.1 2D, and Periphonic 3D (with Height) Compatible Surround Sound Reproduction," AES 115th Int'l Conv., New York City, USA, October 2003 – preprint #5934.
22. G. Theile, "Natural 5.1 Music Recording Based on Psychoacoustic Principles," Proceedings of the AES 19th Int'l Conf., Schloss Elmau, Germany, 6/2001, rev. 10/2001.
23. www.ambiophonics.org, "AmbioFiles" pages with PanAmbio DSP code, "Quick Reference," & "User Guide"
24. R Miller, "Contrasting ITU 5.1 and Panor-ambiophonic 4.1 Surround Sound Recording Using OCT and Sphere Microphones," Proceedings of AES 112th International Convention, Munich, Germany 2002, preprint #5577.
25. A Farina et al, "Recording Concert Hall Acoustics for Posterity," Proceedings of AES 24th Int'l Conf, Banff 2003
26. Farina, A., "Collecting Concert Hall Impulse Responses",Workshop presentation, AES 116th Int'l Conv, Berlin, May 2004
27. H. Wittek, O. Neumann and M. Schäffler, C. Millet, "Studies on Main and Room Microphone Optimization," Proceedings of AES 19th International Conf., Schloss Elmau, Germany, 2001.
28. J. Wuttke, "General Considerations on Audio Multichannel Recording," Proceedings of AES 19th International Conf., Schloss Elmau, Germany, 2001.
29. M Williams, "Multichannel Sound Recording Practice Using Microphone Arrays," Proceedings of AES 24th Int'l Conf, Banff Canada 2003
30. R. Miller, R Glasgal, A Farina, A Torger, E Armelloni, "Multi-format Comparison Demonstration" (simultaneous recordings in ITU 5.1, Stereo, Ambio 2.0, PanAmbio 4.0, PerAmbio 3D), AES 24th Int'l Conf, Banff, Canada 2003.
31. G. Theile, "On the Naturalness of Two-Channel Stereo Sound." J. Audio Eng. Society, Oct. 1991.
32. Y Tao, A Tew, S Porter, "Study on Head Shape Simplification Using Spherical Harmonics for HRTF Computation at Low Frequencies," AES 114th Int'l Conv., Amsterdam, The Netherlands, 2003, preprint #5787.
33. T Lee, D Jang, K Kang, J Kim, D Jeong, H Hamada, "3D Audio Acquisition and Reproduction System using Multiple Microphones on a Rigid Sphere," AES 116th Int'l Conv, Berlin, May 2004, prepring 6135.
34. "Multichannel surround sound systems and operations" – AES Technical Document ESTD1001.0.01-05
35. Miller, R., "Two VLF [SW & Bass Management] Channels?", Workshop presentation, AES 116th Int'l Conv, Berlin, May 2004

Author

Robin Miller is a member of AES and SMPTE, musician and orchestrator, engineer and teacher, and filmmaker recognized by 53 awards including The Peabody. He received a BSEE and has done graduate study at Lehigh University and Temple University. He has more than 40 years experience in music recording and mixing 300 films and television specials plus countless live broadcasts.