WHAT IS SURROUND SOUND ?

Glenn Dickins *
Telecommunications Engineering
Australian National University
Canberra ACT 0200

ABSTRACT: Surround sound now appears as a feature on a myriad of consumer audio products. Much confusion exists regarding the function and capabilities of such systems. This article discusses the range of 'surround sound' systems available and explains two fundamental principles of operation. This leads to a discussion of the capabilities and limitations of current systems and likely future trends.

1 INTRODUCTION

The term "Surround Sound" has been used to describe a number of products and systems. The frequent use of this term has watered down its impact. This article describes a range of past and present surround sound systems discussing their technical merit. It addresses the questions - What is "Surround Sound" ? What can it offer consumers, now and in the future?

Some simple systems merely aim to make a difference in the produced sound. The technical details of these so called 'surround sound' systems are often lost in the marketing hype. Many 'surround' or stereo expanders are little more than selective gain boosters, adding bass, treble and amplifying the stereo difference.

Other surround sound systems offer a complete new approach to the recording and reproduction of audio. They can create a listening experience well beyond that of standard two channel stereo – fully immersive audio. Several multi-channel audio systems and media formats have been developed and are now readily available at a consumer level.

Adoption of new technology requires a reason to change. In the past, advances beyond two-channel stereo, such as quadraphonic, have not had great consumer success. The reader may ask, what has changed now ?

- Consumer surround sound is being driven by 'home theatre' – recreating the movie sound track experience in the home.
- Advanced media such as the Digital Versatile Disc (DVD) has made multi-channel audio practical and convenient.
- The availability and capability of advanced digital signal processing has made the technology for recording and playback of surround audio more affordable.

The range of surround sound product on the market is growing rapidly and will continue to do so.

2 TYPES OF SYSTEM

Acoustical events generate sound pressure waves or variations at the point of observation. This can be described as a sound field – pressure variations about the mean air pressure. A listener at the observation point perceives this sound field through the senses (hearing, touch and perhaps sight through the effect on objects around the listener). To simulate or reproduce an acoustical event we can either attempt to recreate the sound field or just the effect on the senses.

2.1 Sound Field

A sound field system aims to recreate the acoustical pressure field over a region of space. Such a system is independent of the listener in the sense that a representation of an acoustical event is recreated even if no listener is present.

Sound field systems typically use a multitude of speakers to reproduce sounds originating from different directions. The goal may not always be to recreate the exact sound field, but rather something which will be perceived as similar or convey the appropriate information.

This type of system is well suited to delivering surround sound to a large audience. However to accurately represent a sound field over a large region of space a very large number of channels would be required [1].

2.2 Binaural

Rather than recreating an actual sound field, binaural systems attempt to reproduce an appropriate stimulation of the listener's auditory senses. By controlling the sound only in the region of the ears (as with headphones) it is possible to control the excitation of the listener's eardrums. This can result in the perception of complex spatial sound.

Binaural systems deal with only two channels of information. Through direct recording or appropriate processing, these two signals are constructed to represent the pressure variations experienced at the eardrums of a head placed in the desired sound field. When using headphones, the transmission path from the headphone driver to the eardrum can be measured and inverted to compensate for the effects of the ear canal.

Complexities arise with binaural systems due to the fact that each individual has a unique head related transfer function (HRTF) [2]. Psycho-acoustic effects can strongly bias the perceived sound when it is not consistent with visual cues. Despite these drawbacks it is still possible to control the perceived position of sound sources using binaural techniques.

* Also : Lake DSP Pty Ltd
Suite 502, 51 Mountain St
ULTIMO NSW 2007

Binaural systems have the advantage of a smaller signal space. Less information is required to produce a surround sound experience. However, once a signal is reduced to a binaural form, the listener is no longer free to experience the effect of moving in the sound field. Small rotations of the head are useful for resolving sound source locations – this mechanism is lost once a signal is reduced to static binaural form.

2.3 Hybrid Systems

The distinction between the two types of systems is not always easy. Hybrid systems may use a combination of the two.

The premise of a transaural system is to deliver a binaural signal using speakers [3, 4]. Processing of the binaural signals must be carried out to cancel the speaker cross talk and head related transfer functions. The goal is to achieve independent control of the pressure at each of the listener's eardrums using two or more speakers. There is a fundamental limitation of the region of control and the bandwidth. For full bandwidth, accurate transaural delivery the listener and speaker locations must be known to within a few centimetres.

A sound field audio signal can be monitored over headphones. By simulating the transfer functions of the speakers to the listener's ears, an appropriate binaural signal can be constructed from the sound field recording. The transfer functions model the interactions of the playback speakers, listening room acoustics and head related transfer functions.

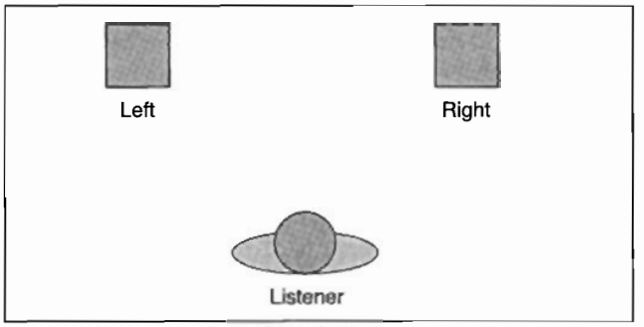

Figure 1. Two channel stereo configuration

3 CURRENT SYSTEMS

3.1 Stereo

Stereo is not confined to two channels [4] – cinema stereo in the 1950's employed four channels of audio [5]. The LP phonograph became the first major consumer stereo format and provided only two channels due to physical constraints. Since then the term stereo has become associated with two channel audio recording and playback. In its basic form a stereo system is a sound field type system which can create the illusion of virtual sound sources between the two speakers (Figure 1).

Depending on the recording techniques used, a stereo recording can contain binaurally encoded information and produce a larger sound image space over headphones or speakers.

The optimal listening position is along a line bisecting the two speakers. This area is known as the 'sweet spot'. As the listener moves away from this region, the quality of the sound field imaging will degrade – sound images initially between the speakers will become unstable and collapse to the nearest speaker.

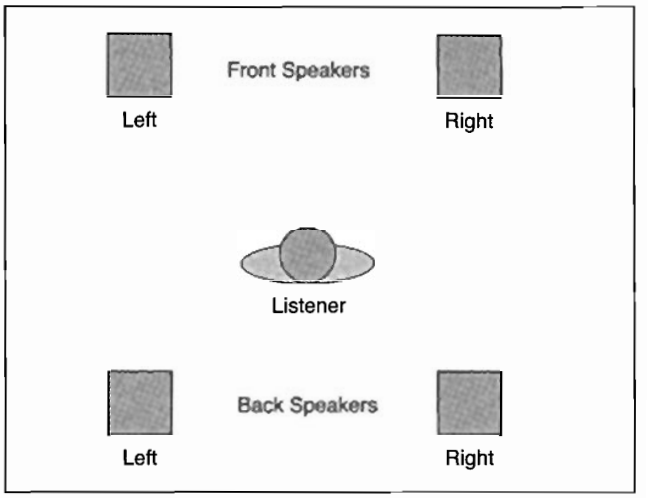

Figure 2. Quadraphonic configuration

3.2 QUADRAPHONIC

Several techniques were developed for delivering more than two channels of audio in the early 70s. Typically the additional two channels were encoded on top of the existing two channels. There were several standards, each requiring its own decoder.

Quad systems represent a sound field type surround system covering 360 degrees in a planar array (Figure 2). The 90-degree separation between speakers tends to leave holes in the sound imaging, particularly at the front centre.

Quad systems were not associated with movie program material and did not offer a significant advantage to consumers for audio program material. The presence of competing and incompatible standards created confusion in the market. As a result of this and other contributing factors, quad systems were not commercially successful.

3.3 Matrix Stereo

In the mid to late 70s, matrix techniques similar to those used for quadraphonic were used for film sound [5]. Four channels of audio information could be encoded on only two channels by exploiting the phase relationships between the channels.

The Dolby MP Matrix encodes two extra channels as matched phase and opposing phase additions to the left and right channels (Figure 3 [6]). In decoding, the centre channel is derived from the sum of the Lt and Rt channels while the surround is derived from the difference. The matrix encoding and decoding process is degenerative and can only achieve 3dB channel separation between adjacent channels [6].

A speaker layout suitable for film sound was adopted with left and right plus a centre dialog channel and a surround effects channel (Figure 4). Dolby Stereo became the standard distribution format for film sound using optical encoding of the Dolby MP Matrix two-channel soundtrack on the film [5].

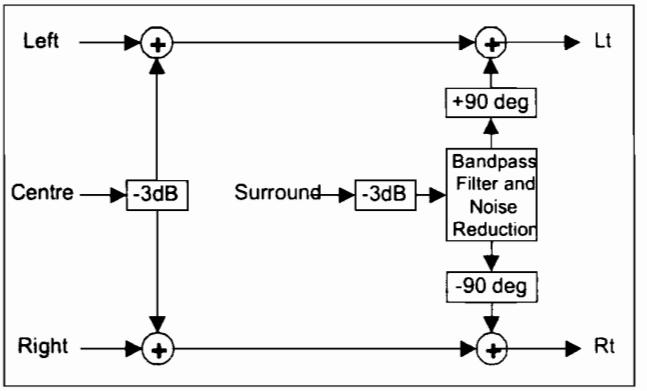

Figure 3. Dolby MP Matrix Encoder

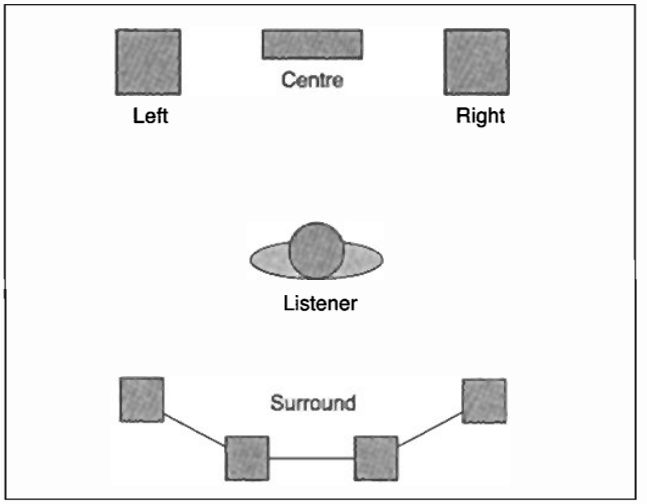

Figure 4. Dolby Surround configuration

Rather than creating an accurate sound field, these systems are concerned with reproducing a desired cinematic experience. The surround channel is used to create enveloping effects and ambience rather than exact sound source image locations. The centre channel is used to create a stable front sound image across the entire audience – a larger 'sweet spot'. This represents a significant improvement over two-channel stereo.

As video recorders became popular in the 80s, matrix stereo surround found its way into the home. Consumer Dolby Surround decoders allowed recreation of the four channel movie sound track from a two channel video source. A pair of diffuse radiating speakers is generally used to appropriately reproduce the surround channel.

Active steering was introduced to improve the decoding process. Through monitoring the relative levels of the decoded channels, the intended sound direction can be estimated. An additional gain stage is included in the decoder to 'steer' the sound in that direction, emphasizing the effective channel separation.

Dolby ProLogic incorporates active steering and became available in consumer products in 1987. This gave an effective channel separation for a steered sound source of 37dB [6]. The additional processing introduces artifacts as the active matrix responds to the input and steers the output – this is known as 'pumping'.

Even with steering, all channel matrixing processes, including Dolby ProLogic, suffer a fundamental limitation. The four channels derived from the matrix stereo signal are not fully independent. Both the surround and centre channels have reduced bandwidth [6]. Complicated acoustical events with many simultaneous dispersed sound sources cannot be accurately recreated.

Another matrixed surround format is Circle Surround. This process claims improved channel separation and steering over Dolby ProLogic [12].

3.4 Ambisonics

Ambisonics is a hierarchical system for the optimal representation of sound fields. It is based on spherical harmonic expansions of the wave equation [1, 7]. The techniques involved have been developed from a combination of acoustical control and psycho-acoustic principles to correctly match important auditory localisation cues.

Ambisonics is not tied to any particular speaker layout and represents a very flexible and generic way of recording, simulating and re-creating three dimensional sound. It has found applications in large audio displays for virtual reality, artistic presentations and theme park entertainment.

Several consumer surround decoders offer an Ambisonic decode mode. A limited amount of program material is available for these decoders in a matrixed stereo format. Although the system is well founded in theory and can deliver excellent results, in practice it has not found a market beyond surround enthusiasts. The process of correctly setting up an Ambisonic decoder and speaker array is quite complex. Hopefully support for this format will continue – it represents a good option for true sound field recordings on emerging multi-channel audio media.

3.5 Multi-channel Digital Formats

To overcome the limitations of matrix surround for movie soundtracks, multi-channel digital formats were introduced. Dolby Digital 5.1 (AC-3) and DTS were introduced as competing standards for theatre sound around 1992. Both use compression algorithms to reduce the multi-channel sound track to a manageable amount of digital data [8, 11, 13]. Dolby Digital typically uses a higher compression ratio than DTS. Qualitative comparison of the two compression schemes is a contentious issue.

Dolby Digital and DTS offer 6 independent channels of audio for a cinema configuration – three front channels, two rear or surround channels and a low frequency effect (LFE) channel (Figure 5). The rear channels are usually line arrays of speakers in a theatre to reproduce a diffuse sound field.

Both of these formats are now available in consumer products. Dolby Digital has secured a greater share of the market by leading the way in available media. More laser discs provide an AC-3 audio stream compared to DTS titles. The option of an AC-3 audio stream has been integrated into the DVD video standard around the world. Dolby Digital transmissions are expected to accompany digital TV broadcasting later this year in the US.

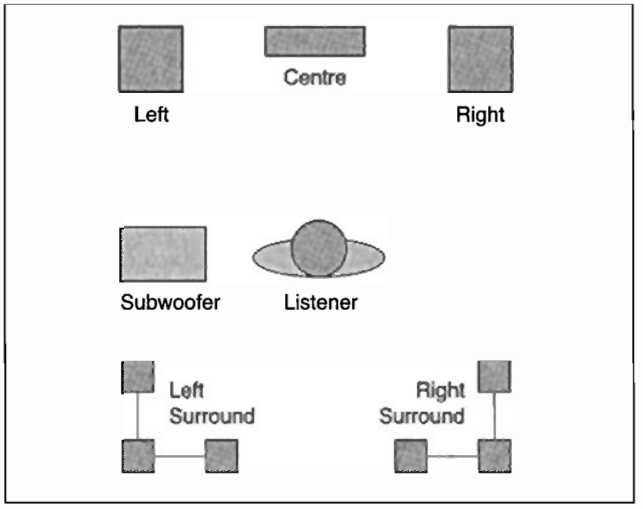

Figure 5. Speaker configuration 5.1

The sound quality is a significant improvement over analog matrix stereo. The overall audio experience can be more convincing and enveloping. With an installed base in place, the 5.1 configuration is also being used for audio only material. There are many DTS recordings of popular music now available on CD [13].

Movie material in the home is brought to life by multi-channel digital audio. As directors and sound engineers are learning to take full advantage of the format, new release movies are delivering powerful and interesting surround sound.

3.6 Sound Enhancements

Many surround products and decoders offer proprietary modes for surround enhancement. Synthesizing additional surround channels can enhance conventional stereo recordings. Although such systems may produce a 'surround sound' effect, it is the result of a processing algorithm rather than the intention of the recording artist. While being a useful selling feature they are often fatiguing to listen to for extended periods.

For source material already in a surround sound format, the unit may add additional ambience or acoustical presence. Movie sound tracks are intended for playback in the environment of a large theatre. To recreate this sound in the home theatre, the system must be accurately set up and calibrated. Additional processing can be used to compensate for the difference in frequency response and acoustical properties of the smaller home theatre. The LucasFilm THX company offers a range of technology and certification of equipment for this purpose [14].

Advances in signal processing algorithms and hardware will continue to provide greater scope for sound enhancements. Consumer products will adopt these as product features and differentiators.

3.7 Virtual Surround Sound

Multi-channel surround formats raise an issue of convenience for consumers. Setting up a full speaker array can be costly and impractical in domestic situations. Some interesting technologies are now emerging in consumer products to address this.

The term 'virtual surround' is being used to describe products that create the impression of a surround sound system using headphones or two speakers [9]. The surround speaker feeds, which would normally be fed to a complete speaker array, are processed to add spatial information.

Over headphones a binaural signal can create the impression of virtual speakers around the listener. Binaural impulse responses can be convolved with the speaker feed signals adding directionality and spatial characteristics [15]. For headphone listening, binaurally processed virtual surround can be vastly superior to a simple 2 channel down mix.

Using two speakers, specially designed filters can be used to create the illusion of sound image locations beyond the span of the speakers. These filters consist of the head related transfer function of the desired virtual sound source position combined with the inverse of the head related transfer function of actual speaker. The net effect is to use the physical speakers to excite the listener's eardrums with a signal approximating that which would be experienced had the sound source been elsewhere (the virtual source location).

The rear surround channels are processed in this way and mixed into the front stereo channels to create the impression of a diffuse sound field to the side or behind the listener. With this type of system, there is a trade off between the quality of the effect and the range of listening locations it will cover ('sweet spot' size). These transaural systems are well suited to an individual listener in a known location [9].

4 FUTURE TRENDS

A significant force in the development of consumer surround sound is the available media. Home video material is the largest source of surround program material. Currently most of this is in matrix surround format (stereo VHS, cable or direct TV). The number of titles released on DVD will continue to increase and more of these will offer full surround audio tracks.

Without being an official standard, the 5.1 speaker configuration (Figure 5) is becoming the dominant consumer configuration. Consumer amplifiers supporting full surround decoding and speaker outputs will become more affordable. With a growing installed base of surround sound systems, more TV programs and audio only material will support the new multi-channel formats. Computer games and multi-media will also embrace and support this format as the computer becomes integrated into the home-entertainment system.

As media capacity and processing ability increase, the cost of multi-channel audio systems will decrease. It is likely that several audio formats will evolve for even more channels. Already higher order systems are used in custom installations like theme park rides and specialty theatres such as IMAX.

The installation of large speaker arrays will be justified for theatres and public spaces but will not be practical for the domestic market. Even five speakers present a logistics problem in the average home theatre. There will be an increasing trend in technologies for reproducing higher order surround formats over a reduced number of speakers.

High capacity sources, such as the DVD, will contain the individual speaker feeds to allow the use of a full array. In

addition to decoding the surround sound, players and amplifiers will perform processing to make optimal use of the speaker setup the listener is using. Cordless headphones and personal speaker arrays will deliver high quality spatial sound. Personal computer or viewing terminals will use passive video techniques to locate the listener and optimize transaural audio delivery [10].

As the number of channels increases, multi-channel audio and true 3D sound begin to merge. Ultimately audio will be represented as sound events or samples with specified locations and acoustical environments. Intelligent players will render this audio information taking into account the listeners speaker configuration and listening environment. Audio compositions will become virtual audio landscapes in which the listener is free to roam and explore or simply pull up a virtual chair, sit down and listen.

REFERENCES

1. M.A. Gerzon, "Periphony: With Height Sound Reproduction" *J. Audio Eng. Soc.* **21**, 2-10 (1973)
2. H.L. Han, "Measuring a Dummy Head in search of Pinnae Cues" *J. Audio Eng. Soc.* **42**, 15-37 (1994)
3. D.H. Cooper and J.L. Bauck, "Prospects for Transaural Recording" *J. Audio Eng. Soc.* **37**, 3-19 (1989)
4. D.H. Cooper, "Comments on Distinction Between Stereophonic and Binaural Sound" *J. Audio Eng. Soc.* **39**, 261-266 (1991).
5. J. Hull, *Surround Sound Past, Present and Future* Dolby Laboratories Inc, San Francisco, 1997
6. R. Dressler, *Dolby ProLogic Principles of Operation* Dolby Laboratories Inc, San Francisco, 1997
7. M.A. Gerzon, "Practical Periphony: The Reproduction of Full Sphere Sound" Audio Eng. Soc. Preprint **1571** (1980)
8. M. F. Davis, "The AC-3 Multichannel Coder" *J. Audio Eng Soc*, presented at 95th convention 1993.
9. G. Brockhouse, "2 Will Get You 5" *Stereo Review*, August 1998, 59-63
10. C. Kyriakakis, T. Holman, H. Nevin and C von der Malsburg, "Immersive audio for desktop systems" *J. Acoust. Soc. Am.* (Abstracts) **103**, 3026 (1998)
11. Dolby Laboratories Technical Papers http://www.dolby.com/tech
12. Circle Surround http://www.surround.net
13. DTS http://www.dtstech.com
14. LucasFilm THX Home Theatre http://www.thx.com/hthx_home.html
15. Lake DSP Home Page http://www.lakedsp.com