Delivering Spatial Audio

Jim Barbour
Swinburne University of Technology
John Street, Hawthorn, Victoria 3122
Australia
jbarbour@swin.edu.au

Abstract

Many sounds artists are now producing their work in multi-channel digital audio, utilizing powerful audio software in very fast computers, and storing sound files on large hard drives. Some works will be played in public performance directly from computer, where multi-channel sound cards are able to deliver multiple high fidelity audio streams simultaneously. However, for more people to experience multi-channel music and sound art, it must be delivered into a listener's environment through existing or new media channels. This paper explores the different media available from physical discs, including CD and DVD, through broadcast media and portable players to digital delivery via the Internet. The possibilities and limitations of each will be considered, focusing particularly on listener perception of quality and ease of use. New directions in each delivery media will also be discussed.

Data Rates

Audio produced or recorded in a digital audio computer is stored as sound files, with their size dependent on the two orthogonal parameters of sampling rate and resolution. The first consumer standard for digital audio was the audio compact disc, CD audio, where two channels of audio are delivered as a stereo recording, and is produced at a sampling rate of 44,100 samples per second per channel, at a resolution of 16 bits. As can be seen below in Table 1, CD audio produces a data stream at 1,411,200 bits per second.

Audio	Resolution (bits)	Sampling Rate (Hz)	Data Rate (bits / sec)
Mono	16	44,100	705,600 bps
Stereo (CD audio)	16	44,100	1,411,200 bps
Mono	24	44,100	1,058,400 bps
Mono	24	48,000	1,152,000 bps
Mono	24	96,000	2,304,000 bps
Stereo	24	48,000	2,304,000 bps
Stereo	24	96,000	4,608,000 bps
6 channels	24	48,000	6,912,000 bps
6 channels	24	96,000	13,824,000 bps

Table 1: Data rates for different audio formats of resolution, sampling rates and audio channels.

Audio and Data Compression

There are two principal types of compression in active use in the production and delivery of digital audio: dynamic range compression and digital data compression. Dynamic range compression is very important in our perception of sound, and has been in use for many decades to maximize the audibility of sound above any noise produced by recording and delivery media: for example, to improve the perceptive quality of audio played from a vinyl record, where surface noise is high and can mask quiet sounds. Dynamic range compression is still very important when audio is digitally recorded and delivered, but is not the focus of this paper.

Digital data compression utilizes software algorithms to reduce the size of digital data files to store more data on a small storage medium or to pass more data through a delivery medium with limited 'bits per second' capacity. If we consider the specifications of CD audio, recorders and players are capable of delivering a data stream of a little more than 1.411 Mbps. This allows stereo audio recorded at 16 / 44.1, (16 bit resolution at 44,100 Hz sampling rate), to be delivered uncompressed, so the original audio data files are unchanged. For 74 minutes of stereo sound on CD audio, this is a little over 700MB of digital data.

Multi-channel Digital Audio Production

From the first introduction of CD audio, there have been many music producers and listeners who have criticized the perceived quality of digital audio as being less enjoyable to listen to than analogue audio, particularly vinyl records, despite surface noise and other aural artifacts introduced by the vinyl medium. There has been much audio research into new technology[1] to improve the perceptive quality of digital audio, based on the fundamental belief that the means of producing, manipulating, storing and delivering audio is otherwise significantly enhanced by digital audio systems. The two principal improvements that have been adopted by audio producers has been increasing the resolution to 24 bits and increasing the sampling rate to 96,000 Hz (24 / 96). While even greater resolution and higher sampling rates are available and in use, most researchers and critical listeners have accepted that 24 / 96 is acoustically transparent and perceptively identical to the original acoustic sound waves, [1]. However, as is immediately obvious in Table 1, 24 / 96 stereo audio produces a data rate of more than 4.6 Mbps, which is well above the capacity of CD audio. Therefore, while audio producers and artists record and produce at 24 / 96, they must use data compression to convert the digital data of their final stereo mix to a lower quality at 16 / 44.1 to deliver using CD audio. Again, much audio research has focused on the best technology to facilitate the format conversion to CD audio, as it remains one of the most readily accepted and purchased delivery media. However, it is clear that CD audio is not capable of delivering the highest quality audio now being produced.

While composers experimented with multiple sound sources surrounding the audience, audio technology was pushed to its limits to playback multiple recorded sound sources around an audience also. Analogue tape recorders with more than two channels were in use in recording studios from the mid 1960's, and were almost immediately taken out of the studio and used for live performance playback. The history of musical composition and experimentation is rich with examples of multi-channel recording and playback over many decades. The quality of digital audio technology has now passed the point where there are no limitations on the perceived quality of digital audio in the production studio, and few limitations on the number of simultaneous channels of audio that can be produced and played directly from a digital audio computer. However, the major limitation still to be negotiated is the delivery of multi-channel digital audio to listeners in their homes, their cars or on portable devices.

1 For example, audio research published by the Audio Engineering Society, www.aes.org.

Discrete versus Ambient versus Matrixed

At this point, we need to distinguish between different content philosophies and different delivery technologies. In surround sound production, there are two principal approaches to the use of the rear channels. A concert hall performance recorded in 5.1 surround has the direct sound from the front of the hall reproduced through the front three loudspeakers, with the rear loudspeakers reproducing the hall reverberation or ambience. Naturally, there will be some ambience in the front channels also, but in the main this would be described as an ambient recording. Alternatively, a recording with real sources behind the listener, which are reproduced directly through the rear loudspeakers, requires the 5 main channels to have discrete content, and is described as a discrete recording, even though there could be significant ambience in all five channels. Discrete moments occur in many movies, where a sound object moves overhead or around the audience and plays directly from the side or rear, but mostly a Director does not want any discrete sounds attracting attention away from the screen. However, for many music and sound art recordings, sounds can and do come from anywhere around the listener, and so the delivery system must be capable of accurate placement of discrete sounds around and behind the listener. Therefore, it can be seen that a matrix system like Dolby Pro Logic [11] can work for ambient recordings, where only reverberant ambience comes from around or behind a listener, but will not be effective for discrete recordings.

Cinema and DVD movies

The cinema industry has set the first standard for multi-channel sound playback with the adoption of the ITU-R BS.775 recommendation for 5.1 surround sound, [2]. This places loudspeakers at the front left, front centre and front right, along with rear left and rear right positions and a subwoofer loudspeaker to reproduce low frequencies effects, the LFE channel. The recommendation requires the five main channels to be full bandwidth audio channels, 20-20,000Hz, while the 0.1 LFE channel is restricted to below 200Hz only. In a playback environment, where the five main channels are unable to reproduce low frequencies due to their physical size, bass management should redirect those low frequencies to the LFE channel. While originally conceived for cinema, 5.1 surround sound has become the audio standard for the new generation of delivery media, the digital versatile disc (DVD), and has been part of the revolution in cinema delivery with the DVD-video format exploding across the world as a commercially viable medium. The 5.1 surround sound format has also been embraced as an audio-only delivery format due primarily to its widespread acceptance by listeners purchasing DVD-video players, many of whom also purchase the audio technology to reproduce surround sound in their homes. However, DVD-video with 5.1 surround sound is not ideal for the delivery of spatial audio as produced by musicians, composers and sound artists for a variety of reasons, with one principal reason being its limitation of horizontal only reproduction.

Surround Sound Data Compression

Music composers, producers and sound artists working with digital audio software are producing spatial audio in 5.1 format at up to 24 / 96 data rates. From Table 1, we know that the final 5.1 surround mix, which occupies six channels of audio data, will require a delivery data stream of approximately 13.8 Mbps at 24 / 96. There will be a very small reduction if the LFE channel is limited to below 200Hz. The maximum data rate from the specifications for the DVD format for all information playing from the disc is 10.08 Mbps, [3]. On a DVD-video disc, the data rate required for the delivery of the video data stream varies from 4.5 Mbps upwards, which means that the audio data stream must be reduced to less than 5 Mbps. This requires significant audio data compression, which reduces audio quality along with size. The first audio data compression formats adopted as part of the DVD-video specification were Dolby Digital AC-3 [11] and DTS Coherent Acoustics [12]. An uncompressed stereo audio soundtrack is also in the specification, and viewers have a choice through menu navigation of which soundtrack they wish to listen to. It is not possible to stream more than one soundtrack simultaneously.

AC-3 and DTS

Dolby Digital AC-3 was originally released in 1992 for the surround soundtrack on a 35mm film, where the digital data is stored on the film itself, utilizing the spaces between the sprocket holes. Due to the space available and the speed of the film through the projector, the data rate available is very restricted, and AC-3 can produce a maximum data rate of 640,000 bps. Therefore, six channels of 24 / 96 digital audio (13.8 Mbps) is data compressed to 640 kbps, a compression ratio of approximately 22:1! When AC-3 is used on DVD-video discs, the data rate is typically 448 kbps, a compression ratio of over 30: 1! When the DVD-video is played, the AC-3 data stream is output as a digital data stream from either an optical or digital coaxial socket and is decoded in an audio amplifier back to the 5.1 loudspeaker feeds. DTS was also released for surround sound in the cinema in 1993, but uses time code on the film to synchronously drive a separate CD-ROM, where the audio is recorded as compressed data. With the data rate limitations of the CD-ROM player, DTS compresses the 5.1 source files to a maximum of 1.536 Mbps, a compression ratio from 24 / 96 of 9:1. Both AC-3 and DTS limit the LFE channel to less than 200 Hertz. The most important question becomes: what is the perceptual quality of the AC-3 and DTS formats compared to the original uncompressed audio? This paper will briefly consider this question in a later section.

Physical Distribution Media for Surround Sound: CD audio

Both AC-3 and DTS can be encoded to a stereo audio file at 16 / 44.1 in the .wav format, which can be burnt to CD audio. The maximum data rates are AC-3: 640 kbps and DTS: 1.284 Mbps. When played on a CD audio player, the digital data stream is output from the optical socket and decoded in an amplifier. Using this format, 5.1 surround sound on CD audio can be manufactured and distributed easily and there have been many commercial releases, particularly with DTS encoding. Data management on the disc also allows an uncompressed stereo data stream to be stored on the disc, and if the optical output is used, surround sound is decoded, while if the analogue outputs are used, standard analogue stereo is produced. Note that only one stream is available at a time, so the maximum number of channels is 5.1 using AC-3 or DTS. Care must be taken during production and with CD audio disc management, as the encoded wav file is unlistenable data hash, which can damage loudspeakers if played at a high volume from the analogue outputs. Great care must be used to identify the content on this encoded disc to a purchaser and the equipment requirements for surround sound playback. Note also there is no capacity for any images. AC-3 and DTS encoders are often included in output options on software programs including ProTools [4], Nuendo [5] and Cubase [5], though usually for a considerable fee added to the software purchase price. Stand alone encoders from Minnetonka Audio [6] are also available.

DVD-video
The DVD-video specification limits the data rates available for audio to less than 6.144 Mbps, with a maximum of eight data streams [3]. These can be used for up to eight uncompressed audio channels, encoded data streams using AC-3, DTS or MPEG, and different language versions. Menu navigation selects the data stream to be used, and decoding can occur either on-board the DVD player or in an external amplifier. The most common surround encoding is AC-3 at 448 kbps, with DTS at 1.536 Mbps as an option, along with an uncompressed stereo soundtrack. DVD-video has been used by composers and sound artists to release audio-only discs, though in many cases they have included still images and music video clips. While it is possible to listen to these discs with the associated television turned off, it is usually necessary to use the on-screen navigation menus to select the desired audio output. The maximum data rate is not sufficient to pass six channels of uncompressed 24 / 96 audio, which is required for the highest fidelity audio reproduction, so DVD-audio has been developed. DVD-video encoders are readily available, including Final Cut Pro [7] and Adobe Premiere / Encore [8], although the options available for audio are often limited or an expensive addition.

DVD-audio
The DVD-audio specification has a maximum peak data rate of 9.6 Mbps, which is still not high enough for 6 channels of uncompressed 24 / 96 audio. Consequently, Meridian Lossless Packing (MLP) [23] is used to reduce the data rate to 7.56 Mbps on average, allowing bit perfect reproduction from DVD-audio of six full bandwidth channels, (it's packed, not compressed!). Note here that the LFE channel is not band limited to below 200Hz, which allows the potential for DVD-audio to carry different forms of spatial audio arrangements other than 5.1, which expands the options available to a sound artist. Unfortunately though, the software required to produce full function DVD-audio discs is very expensive, due to the complexity of options available and the fee payable for MLP. However, there is an inexpensive option with Minnetonka Audio's *Bronze* software [6], which does not have MLP or menu construction, but does allow a maximum data rate for six channels of uncompressed 24 / 48 audio, at 6.912 Mbps and up to 99 tracks within the single menu. The perceptual difference between 24 / 96 and 24 / 48 is virtually inaudible according to research [1], certainly for the average listener and even for many trained professional listeners. There are also now available inexpensive, universal audio players which will playback all current audio formats, including CD audio, DVD-video, DVD-audio, and SACD. Therefore, DVD-audio represents a viable, accessible format for the release of high quality, multi-channel audio recordings.

SACD
The Super Audio CD format released by Sony and Phillips represents a significantly different technical format for audio recording and reproduction [9], which does not use the linear Pulse Code Modulation method of sampling and resolution that is the basis of CD audio and DVD. Instead, it uses a very high sampling rate, one bit resolution process (2.8224 Mbps) called Direct Stream Digital (DSD), which at this point in time is a very expensive software and hardware option on a computer audio workstation. The SACD format also provides very secure copyright protection, as it can only be manufactured under license and they are expensive to produce. While both DVD-audio and SACD produce very high quality audio reproduction and both have their professional supporters and detractors, there have been no definitive listening tests conducted that can show a statistical difference in listener perception of one format compared to the other [1]. With universal players now capable of playing both, there is no difference to a consumer as to which format is used for a commercial release. Note that all universal players at this time decode the data streams for DVD-audio and SACD on-board and playback through six analogue outputs, and do not allow a digital data stream to be output from the optical sockets. While the latest models of Pioneer universal players appear to output DSD through their iLink digital output, the data stream is encrypted and can only be decoded in a Pioneer amplifier. Also, most universal players decode AC-3 and DTS on-board, though both of these formats are available at digital optical outputs.

Broadcasting in Surround
The two main broadcast systems are radio and television, and there are now increasing opportunities for surround sound to be broadcast on both media. High Definition Digital television (HDTV) in Australia and overseas is broadcast with Dolby Digital AC-3 surround sound at 448 kbps [13]. Depending on the source, this could be the 5.1 surround soundtrack of a movie or documentary production, the 5.1 or Pro Logic soundtrack to advertising commercials, or the stereo or mono soundtrack of legacy materials. Standard Definition and analogue TV only has matrix surround broadcast in stereo, usually

Dolby Pro Logic [11], where source material recorded in four channels, left, centre, right and mono surround is combined using a matrix processor into two channels, called Left Total (LT) and Right Total (RT). There is also a move towards broadcasting Digital Radio in surround, using matrix stereo in formats including Windows Media [16], mp3s (a surround version of mp3) [17] and Neural Audio [18]. Space does not allow a greater look at broadcast surround options, but information can be found at the website of Digital Broadcasting Australia [13], Eureka 147 DAB technology for radio [14] and HDRadio by Ibiquity[15].

Internet Delivery for Surround Sound

As a delivery medium, the Internet has the capacity to send any form of digital data from any point in the world to any other, provided there is a network connection. The limitation of delivery is the data rate available, which varies according to the type of connection. High speed cable broadband connections are able to deliver over 1 Mbps reliably, as are fibre-optic backbone connections used by Universities and some industries. Home connections with ADSL broadband can reach 1 Mbps, while dial-up connections still hover around 30-40 kbps. Consequently, delivering digital audio data in real-time is only possible if the data rate is less than the receiver's connection allows. While there have been many attempts to establish real-time audio delivery over the Internet using data compression, nobody expects high quality audio to be available in real-time. Data compression formats including Windows Media [16] and mp3 [17] have been developed, among other reasons, to meet the demand for real-time audio over the Internet, and range from very low compression ratios to very heavily compressed options. Most people who listen to real-time audio over the Internet are now accustomed to lower quality audio reproduction. However, the great advantage of Internet delivery is downloading files.

Download rather than Real-time

If we remove the requirement for real-time audio, the Internet is a very powerful tool for delivering audio to listeners and to establish new markets for purchasing sound recordings. While the time to download a file will vary according to the connection, virtually everybody with Internet access can download the highest quality audio file made available by an artist or producer. Therefore a ten minute surround recording in 5.1, uncompressed at 24 / 96 will be just over one GB, and it could be downloaded!! Not realistic, but possible! However, one example of successful Internet delivery is the radio production department of the Swedish Radio[10], the equivalent of the ABC in Sweden. They have been producing radio programs in 5.1 surround for several years now, including drama, music, comedy and sport, and convert them to DTS wav files which can be burnt directly onto CD audio. They have made the DTS files freely available on their website for downloading, and they have had more than four million downloads! The ten minute example above, when encoded to DTS at 1.284 Mbps, occupies about 96 MB, which would take a long time on a dial-up modem, but is relatively quick on cable broadband. Other sound artists and producers have followed this example and have made high quality surround recordings readily available on-line. Of course, the on-line trade in music using low quality audio codecs, like mp3, has undermined the potential for significant sales of music over the Internet, though there are now well established sites where music can be purchased. Therefore, the Internet is a great service for downloading music and will become even more useful and popular for high quality music delivery as data rates improve and broadband connections are more widely adopted.

Portable hard drives and flash memory sticks

The rapid uptake of portable hard disc and flash memory technologies introduces a new possibility for surround music delivery. The rapid acceptance by consumers of the iPod and other portable hard drive and flash memory recorders and players has led another revolution in technology. Currently, all players are mono or stereo, and mainly use the mp3 format for data compression. However, most players will connect to a computer through a USB interface, which allows the use of the player as a portable memory device. It is possible to store surround sound formats on these devices as data files, either as uncompressed wav files or any of the formats discussed in this paper. Consequently, the next stage is to connect these devices directly to a playback amplifier with decoding software and loudspeakers, to allow surround sound playback. The marketing of home computers as Home Media Centres takes this a step closer to realization, connecting hard drives and portable memory devices directly to video display units, amplifiers and loudspeakers. With surround sound music available for downloading via the Internet and portable devices capable of storing and transporting the files, there may be no need soon for any CD or DVD media. However, there will still be a need for data compression, to facilitate more rapid downloading and to store more material on portable devices.

Low Bit Rate Encoders
There is currently substantial development work by several audio coding technology companies trying to reduce the data rates for acceptable discrete surround sound delivery to below 192 kbps. Some of these formats include Windows Media [16], mp3s (a surround version of mp3) [17] and Neural Audio [18], with proponents suggesting they can use rates as low as 64 kbps, which represents a compression ratio in excess of 200:1 from 24 / 96! If the producers of high quality stereo recordings at 24 / 96 are upset when their work is converted to mp3 at low bit rates, how will they feel when their surround recordings are reduced to even 192 kbps in mp3s? Will radio in surround be a major sales point-of-difference if the quality of the service is compromised by low bit rates? What styles of music might best survive these low rates? Will ordinary listeners perceive a difference anyway?

Subjective Consumer Evaluation of Multi-channel Audio Codecs
The author is currently conducting research into the question: Can ordinary listeners using standard consumer technology perceive any differences between original, uncompressed 24 / 48 sound recordings and any of the multi-channel audio codecs discussed in this paper: AC-3, DTS, WMA and mp3s? Source material includes both discrete and ambient recordings, ranging from Wagner to Linkin Park, and also audience applause. Data compression rates vary from DTS at 1.284 Mbps to mp3s at 192 kbps and WMA at 128kbps. Preliminary results indicate that AC-3, DTS and WMA at 440 kbps cannot usually be distinguished from the uncompressed original. However, WMA and mp3s at 192 kbps have perceptibly audible differences especially for discrete rather than ambient content. A paper with full details is in preparation.

Alternative Spatial Coding Formats
5.1 surround sound has many limitations when attempting to deliver an accurate spatial impression of a recorded or created soundfield, including poor localization at the sides and rear and no elevation. Its strength is its installed listener base through DVD-video systems and a substantial content catalogue. There have been attempts to create alternative techniques for delivering a more accurate representation of spatial audio, including Ambisonic and 5.1 with height. These two will be considered here with details of delivery techniques and reproduction requirements, and a consideration of the possible results.

Ambisonic Production, Delivery, Reproduction
The Ambisonic concept is an extension of the original Blumlein stereo pair technique [20], and involves measuring a soundfield at a single point in space as four component mono recordings. The W component represents an omni-directional recording at this point. The X component is a bi-directional recording with positive phase facing zero degrees forward, and negative phase at 180 degrees behind. This recording has minimum pick-up from the left-right and above-below. The Y component is a bi-directional recording with positive phase facing left at 90 degrees, with minimum pick-up from front-back and above-below. The Z component is a bi-directional recording with positive phase directly up, with minimum pick-up from front-rear and left-right. The diagram below shows these four components.

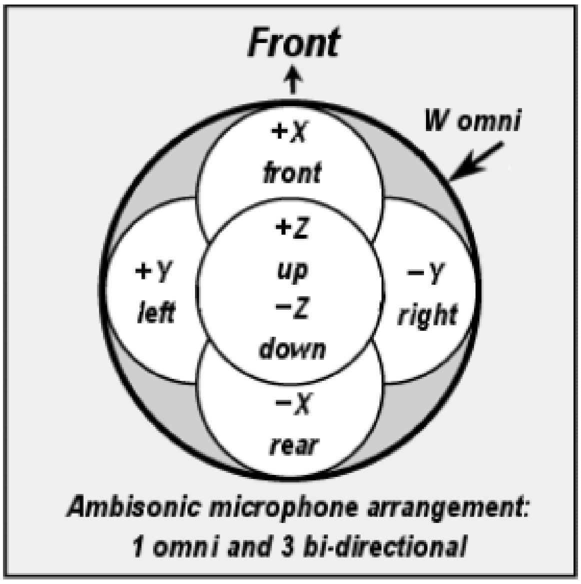

Figure 2: Ambisonic microphone arrangement

Auditory Scene Analysis
Thus, any sound source at any location in space will have its acoustic output recorded by the W, X, Y and Z microphones in unique proportions, depending on its location relative to the point of recording. One version of the combined microphones is called the Soundfield Microphone™ [21], and the output is four analogue mono signals. Using auditory scene analysis, we can break up any sound environment into the total number of individual sound streams, each generated by a sound source. The Soundfield Microphone will simultaneously encode each sound stream into its WXYZ components and the total output from the microphone will be an encoded sound field, called B-format. To recreate the original soundfield, we specify the location of all loudspeakers used for reproduction as XYZ directions, and the decoder will then assign

the correct proportions of the recorded B-format signals to each loudspeaker. For a more detailed explanation of what is a very complex process, please consult the Ambisonic website [22].

B-format and G-format

Interest in Ambisonic theory and practice has been regenerated in the last few years by the introduction and uptake of 5.1 surround sound, as there is now greater perception of spatial audio among listeners, and producers and artists now have an accepted universal format to work with. New digital audio technology has further advanced Ambisonics, as the very high quality of audio recordings now enhances the soundfield recreation and more powerful computers make it possible to manipulate the B-format soundfield in new and interesting ways. However, while B-format recordings are easily decoded in audio software and reproduced in a studio, there are currently only a few very expensive, stand-alone amplifiers capable of decoding B-format and adjusting the outputs for different loudspeaker arrangements [23]. Another disadvantage requires many loudspeakers to accurately reproduce the soundfield from a B-format recording, typically using six loudspeakers at floor level arranged equally around the centre position at 60 degree intervals, and another six at ceiling height, also at 60 degree intervals [22]. This large number of loudspeakers and their arrangement is well beyond the typical home user! However, it is possible to convert B-format recordings to the 5.1 layout, called G-format by Ambisonic users, which loses many of the benefits of Ambisonics but does deliver into the current preferred surround format. Delivery of Ambisonic B-format requires four full range audio channels, which is certainly within the potential of uncompressed DVD-audio, and the B-format components survive encoding with AC-3 or DTS, provided the decoded outputs are then Ambisonically decoded to the reproduction environment. B-format is also an excellent format for Internet delivery and downloading.

The Loudspeaker as a Source

One significant problem encountered by many producers and artists is the tendency for reproduced sounds to appear to come from one loudspeaker only, sometime referred to as speaker detent. This is more noticeable when listening slightly away from the sweet spot, at the very centre of the loudspeaker array. The panning algorithms in most audio software move a sound between two loudspeakers, eg. left to centre or left front to left rear, with an amplitude only operation, with no change in time delay or phase relationships between the two positions. While this can be stable at the sweet spot, off-centre listening introduces the precedence effect with the sound appearing to come from the closest loudspeaker. One advantage of Ambisonic processing is that every position is a combination of four components, and the reproduction of every sound requires some output from every loudspeaker. Therefore, panning using B-format principles reduces loudspeaker detent, creating a more homogeneous soundfield with the loudspeakers less evident. This holds true for a G-format final mix also. To aid this production, there are VST plug-ins available for B-format encoding and decoding, and a B-panner is also available to allow experimentation with Ambisonic panning [24].

Adding the Height Dimension

In a paper delivered by this author at ACMC2004 in Wellington [25], the advantages of adding height to audio recording and reproduction were canvassed with possible recording and reproduction techniques explained. In principle, the soundfield is broken into three stereo pairs, left-right front, left-right rear and left-right elevated. The resulting six channel recordings are reproduced with the four appropriate loudspeakers from the 5.1 array, not using the centre front or LFE, and the addition of two loudspeakers overhead. Further experimentation has confirmed the aural excitement that can be created using this new array, and a method of delivery has been tested and found satisfactory. The six channels are delivered using DVD-audio, with the arrangements as follows: left-right front as normal, left-right rear as normal, left elevated as centre channel and right elevated as LFE channel. As discussed previously, the six channels on DVD-audio are uncompressed and full bandwidth, particularly the LFE channel which is not band limited to below 200Hz. Therefore the LFE channel is capable of delivering the right elevated audio unaffected. On replay, the DVD-audio is decoded on-board a universal player, and the analogue outputs from centre and LFE are patched to the elevated loudspeakers. Where a listener does not have overhead loudspeakers, but instead plays the disc through a 5.1 array, the left overhead channel plays through the centre front loudspeaker and the right elevated channel plays through the subwoofer. Where the recording is of ambient nature, reproduction is virtually identical in affect. When the recording is discrete, there can be some unusual sound movements when sounds travel through the front centre position, but for experimental music it is rarely noticed or commented upon by listeners. Note also that an Ambisonic B-format recording is reproduced very well over the elevated array, and brings out some of the height components that are missing from the G-format reproduction.

Conclusion

Delivering spatial audio requires media capable of storing large files and providing a data stream suitable for the accurate recreation of a three-dimensional soundfield. The playback of high quality, uncompressed multi-channel digital audio requires very high data rates, which must be reduced using data compression so that they can be delivered using media channels currently available that have high consumer acceptance, including CD audio, DVD-video, DVD-audio and SACD. HD Digital Television in Australia can deliver good quality spatial audio, while SDTV and analogue television are not providing a high quality, discrete surround sound experience. New radio broadcasting services including DAB and HDRadio use extreme data compression which compromises the quality of spatial audio. The results of subjective testing indicate that data compression can provide good spatial audio provided the data bit rate is sufficiently high. Alternative spatial audio formats including Ambisonics B-format and 5.1 with height offer solutions to full three-dimensional reproduction, and DVD-audio software and players represents a viable, accessible format for the release of high quality, multi-channel audio recordings. The Internet can be used for delivering spatial audio using high quality formats, providing the files are downloaded. New portable data memory technology allows the storage and playback of audio files, which represents an exciting way forward for consumer access to high quality spatial audio.

References

[1] Stuart, JR, 2004, *Coding for High-Resolution Audio Systems*, Journal of the Audio Engineering Society, volume 52, number 3, March 2004.

[2] ITU-R BS.775, 1993, *Multi-channel stereophonic sound system with or without accompanying picture*. International Telecommunications Union, Switzerland.

[3] Holman, T, 2000, *5.1 Surround Sound, Up and Running,* Boston, Focal Press

[4] Pro Tools software and hardware: www.digidesign.com

[5] Nuendo and Cubase software: www.steinberg.net

[6] Minnetonka Audio software, including DVD-A Bronze, AC-3 and DTS encoders: www.minnetonkaaudio.com

[7] Final Cut Pro video editing software: www.apple.com/finalcutpro/

[8] Adobe Premiere and Encore DVD software: www.adobe.com

[9] SACD information: www.sonymusic.com/sacd/history

[10] Swedish Radio: http://www.sr.se/multikanal/english/e_index.stm

[11] Dolby Laboratories Inc.: Dolby Pro Logic, Dolby Digital: www.dolby.com

[12] DTS Coherent Acoustics: www.dtsonline.com

[13] Information on HDTV in Australia and overseas: www.dba.org.au/ and www.digitaltv.com.au/index.html

[14] Digital Audio Broadcasting using Eureka 147: http://www.worlddab.org/eureka.aspx

[15] Ibiquity HD Radio: www.ibiquity.com/hdradio/

[16] Windows Media encoders and decoders: www.microsoft.com

[17] mp3 surround: www.iis.fraunhofer.de/amm/download/mp3surround/index.html

[18] Neural Audio: www.neuralaudio.com/

[19] ITU-R BS.1534, 1994-7, Methods for the Subjective Assessment of Small Impairments in Audio Systems including Multichannel Sound Systems, International Telecommunications Union, Switzerland.

[20] Blumlein reference: www.er.uqam.ca/nobel/k24305/basic_ambisonics.html

[21] Soundfield microphone, www.soundfield.com

[22] Ambisonics website, www.ambisonic.net

[23] Meridian Audio: MLP and Ambisonic decoding amp: www.meridian-audio.com/p_861.htm

[24] VST plugin development detailed at www.ambisonics.net

[25] *Exploration of the Height Dimension in Audio Reproduction*, J. Barbour, ACMC 2004, Wellington, NZ