Two new consumer audio disc formats are described with regard to their principal audio features. Issues of player and disc compatibility are discussed and it is concluded that while it may be technically feasible to manufacture compatible software or hardware, commercial considerations will decide whether this happens in practice.

0 INTRODUCTION

The purpose of this paper is to summarise the principal features of two new consumer audio disc formats – the DVD-Audio (DVD-A) and the Super Audio CD (SACD). While they have much in common, they also have important differences that make them partially incompatible. This paper will concern itself mainly with the audio features of the discs as opposed to the transport stream or optional text, graphics and video features. The information presented here is largely based upon information provided by the DVD-Audio working group and Sony/Philips (see acknowledgments). Consequently there are few formal references. Potential licensees or users of this technology are advised to consult the appropriate standard ‘books’ that specify these formats in detail.

1 NEW CONSUMER AUDIO DISC FORMATS IN GENERAL

The Compact Disc (CD) has been commercially available for over 15 years. It is now possible to create optical discs with considerably more storage capacity, and with higher data transfer rates than the CD. In recent years optical disc manufacturers have developed a high density format with the same physical dimensions as the CD but which holds many times more data. This is called the DVD (digital versatile disc). The data capacity of the original CD is around 650 Mbytes, whereas the new DVD formats have capacities of a number of Gbytes, depending on the number of layers and sides (the discs can be dual layer, and/or double-sided). A single layer DVD-Video or -Audio disc has a capacity of 4.7 Gbytes. The data transfer rate has a peak value of 9.6 Mbit/s, although this is restricted in certain data objects.

The increased capacity and data transfer rate of this new disc format has many possibilities from an audio standpoint. The main audio features to benefit from these specifications are multichannel surround sound and increased sampling rates/resolutions compared with CD.

2 DVD-AUDIO

Version 1.0 of the the DVD-Audio format specification has been approved by the format’s steering committee. The disc is part of the family of DVD discs that also includes the DVD-Video, DVD-ROM, DVD-R and DVD-RAM. The audio format is the last of these to appear, and has been designed to take into account the recommendations of an International Steering Committee (ISC) that made a series of requirements for a new audio disc format [1].

2.1 Sample frequency and resolution

The format is more versatile in respect of sample frequency than DVD-Video, having also accommodated multiples of the CD sample frequency of 44.1 kHz as options (the DVD-Video format allows only for multiples of 48 kHz). Consequently, the allowed sample frequencies for DVD-Audio are 44.1, 48, 88.2, 96, 176.4, 192 kHz. This allows for audio bandwidths in excess of 20 kHz if record companies consider this important.
The sample frequencies are split into two groups – multiples of 44.1 and multiples of 48 kHz. While it is possible to split frequencies from one group among the audio channels on a DVD-A (see below), one cannot combine frequencies across the groups for reasons of simple clock rate division. Bit resolution can be either 16, 20 or 24 bits per channel, and again this can be divided unequally between the channels, according to the channel group split described below.

2.2 Number of channels and playing time
Playing time depends on the way in which producers decide to use the space available on the disc, but various features combine to make it possible to store at least 74 minutes of stereo audio even at the highest sample rate and resolution (192/24). Other modes are possible, with up to 6 channels of audio playing for at least 74 minutes, using combinations of sample frequency and resolution, together with an optional lossless packing technique known as MLP (see below). 6 channel audio can only operate at the two lower sample rates of either class (44.1/88.2 or 48/96).

A downmixing technique known as SMART (System Managed Audio Resource Technique) is mandatory in players but optional for content producers. It enables a stereo downmix of the multichannel material to be made in the player but under content producer control. The gains, phases and panning of each audio channel can be controlled in the downmix [2]. A separate two-channel mix (L0/R0) can be included within an MLP bitstream. If a separate stereo mix is provided on the disc then this is automatically used instead of the player downmix.

All modes other than mono or 2-channel have the option to split the channels into two groups. Group 1 would normally contain the front channels (at least left and right) of the multichannel balance, while Group 2 could contain the remaining channels. This is known as scalable audio. The resolution of Group 2 channels can be lower than that of Group 1, enabling less important channels to be coded at appropriate resolutions to manage the overall bit budget. The exact point of the split between the channel groups depends on the mode, and there are in fact 21 possible ways of splitting the channels.

It is also possible to 'bit-shift' channels that do not use the full dynamic range of the channel. For example, surround channels that might typically under-record compared with the front channels can be bit shifted upwards so as to occupy only the 16 MSBs of the channel. On replay they are restored to their original gains.

It is not mandatory to use the centre channel on DVD-Audio. Some content producers may prefer to omit a centre speaker feed and rely on the more conventional stereo virtual centre. The merits or demerits of this continue to be debated.

2.3 MLP
Meridian Lossless Packing will be licensed through Dolby, and is a lossless coding technique designed to reduce the data rate of audio signals without compromising sound quality. It is in some ways the 'Trojan Horse' of DVD-Audio as it is mandatory in all players and enables a number of audio features not found in the basic audio stream specification, such as hierarchical channel encoding for surround sound (e.g. Ambisonics), and up to 64 channels of audio data. Individual streams can be flagged for certain speaker feeds if required [3]. It has both a variable bit rate mode and a fixed bit rate mode. The variable mode delivers the optimum compression for storing audio in computer data files, but the fixed mode is important for DVD applications where one must be able to guarantee a certain reduction in peak bit rate.

The use of MLP on DVD-A discs will be optional, but is an important tool in the management of bit budget. Using MLP one would be able to store separate two-channel and multichannel mixes on the same disc, avoiding the need to rely on the semi-automatic downmixing features of DVD players. Owing to the so-called Lossless Matrix technology employed, the L0/R0 downmix that can be made at the MLP mastering stage takes up very little extra space on the disc, owing to redundancy between the multichannel and two channel information. MLP is also the key to obtaining high resolution multichannel audio on all channels without scaling.

2.4 Disc layers
All DVD discs have the possibility to be made either single or dual layer. While most will probably be single layer, it is up to the record companies to decide what versions to issue. Recent developments suggest that DVD-A players will be designed to handle regular CD format as well, and that it will be technically possible to issue dual layer hybrid discs with a high density layer and a CD player. This was originally understood only to be a feature of SACD. The latest manoeuvres in this sensitive area seem to have resulted in the decision about whether to issue hybrid discs being left to the record companies. While it is technically possible to issue many combinations of two layers containing different formats, commercial decisions in the software arena will determine their availability.

2.5 Disc and player types
There will be two types of DVD-Audio disc, one containing only audio objects and the other (the DVD-AudioV) capable of holding video objects as well. The video objects on a DVD-AudioV will be just the same as DVD-Video objects and therefore can contain video clips, Dolby AC-3 compressed audio and other information. In addition, there is the standard DVD-Video disc. There will effectively be three DVD player types on the market (audio, universal and video).
There will therefore be three disc types and three player types that have different amounts of cross compatibility, as shown in Figure 1. DVD-AudioV discs should play back in audio players, universal players and the video objects should play back on video players. The requirement for video objects on DVD-AudioV discs to contain LPCM audio was dropped at the last moment so that such objects could only contain AC-3 audio if desired. This means that an audio disc could contain a multichannel AC-3 audio stream in a video object, enabling it to be played in a video player. This is a good way of ensuring that a multichannel audio disc plays back in as many different types of player as possible, but requires that the content producer makes sure to include the AC-3 video object in addition to MLP or LPCM audio objects. The video object can also contain a DTS audio bitstream if desired.

2.6 Digital outputs and copy protection on players
The first generation of DVD-Audio players is expected only to have standard IEC958 digital outputs at the basic sampling frequency, even if the disc contains higher resolution information. This is because agreement has not been reached at the time of writing over a universal means of copy protection for digital outputs. Future players are expected to be equipped with IEEE1394 (Firewire) interfaces which will enable the full high resolution information to be transferred, possibly including the capability to carry the undecoded MLP bitstream for external decoding into alternative multi-channel forms.

Documentation for MLP claims that the technology is capable of including a watermark in the bit stream, but this would clearly only apply to MLP-encoded bit streams, not to uncompressed LPCM. Further discussions are apparently required before final agreement on universal copy protection for DVD-A can be finalised. So far as can be determined there is no provision for visible and invisible watermarking of the data on the disc, as implemented in SACD (see below).

3 SUPER AUDIO CD (SACD)
Version 1.0 of the SACD specification has now been approved and has been issued to the first licensees as the 'Scarlet Book'. The disc uses high density optical technology for the high density layers, based on DVD principles. Consequently it is understood that the disc can be physically read by the same mechanisms that read DVDs. The data stored on the disc, though, represents audio in a different form to that used on DVD. Instead of using multibit LPCM it uses 1-bit PCM coding known as DSD (Direct Stream Digital) [4]. The disc also addresses the requirements for a new audio disc format set out by the ISC.

3.1 Sample frequency and resolution
The sample frequency used by SACD is 2.8224 MHz.
This extremely high sampling frequency (64 times the basic CD frequency of 44.1 kHz) is possible because of the 1-bit delta-sigma quantising process with noise shaping that shifts quantising noise into the ultrasonic region of the spectrum. Noise within the normal audio band is very low, allowing a dynamic range of approximately 120 dB up to 20 kHz. Early documentation [5] also indicates that a sample frequency of 48 times 44.1 kHz may also be an option, and it is possible (although not confirmed) that this may be necessary for concurrent multichannel mixes. The high sample frequency leads to a usable audio bandwidth of approximately 100 kHz.

3.2 Number of channels and playing time
SACD aims to provide a playing time of at least 74 minutes for both two channel and six channel balances. The disc is divided into two regions, one for two channel audio, the other for multichannel, as shown in Figure 2. It is understood that the first players will be two channel only, but that multichannel devices will be introduced in due course. A lossless data packing method [6] known as Direct Stream Transfer (DST) is intended to be used to achieve roughly 2:1 data reduction of the signal stored on disc so as to enable high quality multichannel audio on the same disc as the two channel mix.

3.3 Disc layers
SACDs can be manufactured as single or dual-layer discs, with the option of the second layer being a Red Book CD layer (the so-called ‘hybrid disc’). Initially it was planned that all SACDs would be hybrid to retain CD compatibility, but the requirement for this seems to have been relaxed. The onus has been placed upon the record companies to decide whether they wish to manufacture hybrid discs or not. SACD players will also be able to play conventional CDs.

3.4 Additional content
SACDs, not being a formal part of the DVD hierarchy of standards (although using some of the optical disc technology), do not have the same options for DVD-Video objects as DVD-Audio. The disc is designed first and foremost as a super high quality audio medium. Nonetheless there is provision for additional data in a separate area of the disc. The content and capacity of this is not specified but could be video clips, text or graphics, for example.

3.5 Watermarking
Sony and Philips have paid considerable attention to copy protection and anti-piracy measures on the disc itself. Comprehensive visible and invisible watermark-
ing are standard features of the SACD. Using a process known as PSP (Pit Signal Processing) the width of the pits cut into the disc surface is modulated in such a fashion as to create a visible image on the surface of the CD layer, if desired by the originator. This provides a visible means of authentication. The invisible watermark is a mandatory feature of the SACD layer and is used to authenticate the disc before it will play on an SACD player. The watermark is needed to decode the data on the disc. Discs without this watermark will simply be rejected by the player. It is apparently not possible to copy this watermark by any known means.

Encryption of digital music content is also optional, at the request of software providers.

3.6 Digital outputs on players
The first players will not have digital outputs carrying DSD data, as a standard interface with copy protection has not yet been agreed. Like DVD-A players, the first generation of SACD players will have standard IEC958 digital outputs. The companies are waiting for industry-wide agreement on a new digital output standard, possibly based on IEEE1394, with suitable copy protection properties.

4 IMPLEMENTATION OF BOTH FORMATS
From a physical point of view the two disc formats differ very little from each other, making it technically feasible to manufacture players that will play any of the new disc formats including all versions of DVD, CD and SACD; single-layer, dual-layer and hybrid construction. It is becoming increasingly clear, therefore, that it will be a commercial rather than a technical decision regarding which options to implement, based on various alignments of the parties involved. Software producers will also have to decide whether to make hybrid discs that have a CD layer, or whether perhaps to make discs with both SACD and DVD layers. The latter would be likely to be prohibitively expensive from a licensing point of view, and the more likely solution is for players to be designed to play both types of disc.

ACKNOWLEDGMENTS
The author would like to thank Bike Suzuki, Chairman of the DVD-Audio working group, and David Walstra of Sony for much of the information on which this paper was based. Responsibility for the interpretation in this paper is, of course, the author’s.

REFERENCES


