Super Audio CD Format

Jan Verbakel
Philips Consumer Electronics Eindhoven The Netherlands

Leon van de Kerkhof
Philips Consumer Electronics Eindhoven The Netherlands

Muneyasu Maeda
Sony Corporation Japan

Yoshizumi Inazawa
Sony Corporation Japan

Presented at the 104th Convention 1998 May 16-19 Amsterdam

AN AUDIO ENGINEERING SOCIETY PREPRINT

Super Audio CD Format

Jan Verbakel, Leon van de Kerkhof

Philips Consumer Electronics, Eindhoven, the Netherlands

Muneyasu Maeda, Yoshizumi Inazawa

Sony Corporation, Tokyo, Japan

The Super Audio CD is the next generation audio disc. It is based on a hybrid disc concept, that provides compatibility with the Compact Disc. The Super Audio CD contains two layers: a CD layer and a high-density layer. The CD layer contains Red Book CD audio. The high-density layer contains both very high quality 2-channel stereo and very high quality multichannel audio. Besides the two types of audio, the high-density layer offers room for supplementary data, synchronised with the audio data, and a separate Extra Data area. The audio on the high-density layer is Direct Stream Digital (DSD) encoded: the audio is represented with a resolution of one bit, at the sampling frequency of 2.8224 MHz. The application of lossless coding, especially developed for DSD audio signals, allows to store 74 minutes of both 2-channel stereo and multichannel audio on the high-density layer.

Introduction

Advances in optical storage techniques allow a significantly higher amount of data to be stored on a 12 cm optical disc than what was possible at the time CD was developed. Several initiatives have been undertaken to define recommendations and requirements for an audio format exploiting this higher data capacity to the benefit of music lovers, artists and recording engineers. Common to these recommendations are an extended frequency range, higher dynamic range and support for multichannel audio. Next to this, the music industry has shown great interest in CD compatibility.
The Super Audio CD fulfils these requirements. It offers a frequency range of up to 100 kHz, a dynamic range of more than 120 dB up to 20 kHz, 2-channel stereo and multichannel audio on one disc. Furthermore, it offers full CD compatibility in the sense that a current Compact Disc can be played on a SACD player *and* a Super Audio Compact Disc can be played on a current CD player. This compatibility is achieved by a hybrid disc concept.

In this paper, the Super Audio CD format is presented in further detail.

The Hybrid Disc

Figure 1 shows the hybrid concept of the Super Audio Compact Disc [1]. The disc consists of two substrates, one covered with a high-density (HD) layer and one covered with a CD layer. The 2 substrates are bonded back to front together, and are read out via the same surface.

The high-density layer has a certain reflectance for a wavelength of 650 nm (the HD laser wavelength) and a high transmission for a wave length of 780 nm (the CD laser wavelength). This layer can be read out by a Super Audio CD player (with a 650 nm laser), but is virtually invisible to the 780 nm laser of a conventional CD player. See also figure 2. The distance of the read-out surface to the high-density layer is 0.6 mm. The storage capacity of the HD layer is 4.7 Gbyte.

The CD layer has a high reflectance at 780 nm, and lies at a distance of 1.2 mm from the read-out surface. This layer is thus perfectly visible by a conventional CD player, reading through the transparent HD layer. The storage capacity of the CD layer is about 750 Mbyte.

The hybrid disc concept ensures compatibility with existing CD players: any regular CD player can playback the CD layer of a Super Audio CD, while new Super Audio CD players can playback the High Density layer and the CD layer. This is depicted in figure 2.

Sector Layout

The data on a Super Audio CD is grouped in sectors of 2064 bytes according to the following structure:

Identification Data (ID)	4 bytes
ID Error Detection (IED)	2 bytes
Reserved	6 bytes
Main Data	2048 bytes
Error Detection Code (EDC)	4 bytes

Error correction and modulation

After scrambling, 16 of these sectors are collected to form a Error Correction Code (ECC) block, providing a powerful error correction scheme based on a Reed Solomon Product Code. The rows of the ECC Blocks are interleaved, Recording Frames are formed, which are modulated with the "EFM-plus" modulation code. This EFM-plus is based on the same basic rules as the EFM in CD, however with an improved efficiency. After this modulation process the data is structured in Physical Sectors which are recorded onto the disc.

CD Application layer

The specifications of the CD layer are fully according to the Red Book standard [2]. Any regular CD player will be able to play this layer. The details are commonly known and will not be discussed here.

HD Application layer

The HD layer contains two audio areas. The inner one stores 2-channel stereo audio, while the second area stores multichannel audio supporting up to 6 channels. The audio content of the two areas is independent.

Structure of the HD layer

The structure of the HD layer is depicted in figure 3. The Information Area of the High Density layer is subdivided in the Lead-in Area, the Data Area and the Lead Out Area. The Data Area is again subdivided in the File System, the Master TOC, the 2-Channel Stereo Area, the Multichannel Area and the Extra Data Area.
The 2-Channel Stereo Area and the Multichannel Area have the same basic structure. Each Audio Area contains an Area TOC and the Audio Tracks. The 2-Channel Stereo Area contains only 2-Channel Stereo Tracks, the Multichannel Area contains only Multichannel Tracks. The optional Extra Data Area may contain data of any kind.

Access methods

The Super Audio CD supports two access methods:

- Using a hierarchical TOC structure, with a Master TOC and two area TOC's
- Using a UDF and/or an ISO 9660 file system

Every disc contains the TOC structure. The use of file systems is optional.

The TOC structure

The TOC has a two level structure. The highest level is the Master TOC, and the lower level is formed by two Area TOC's, one in the 2-Channel Stereo Area and one in the Multichannel Area. The Extra Data Area does not contain an Area TOC.

The Master TOC is stored on three fixed locations, at sectors 510,520 and 530. These positions are chosen such that probability that the Master TOC cannot be read due to a damaged disc is minimal. The Master TOC contains general information on the high density layer and information on the size and location of the Audio Area's. The second and higher Sectors of the Master TOC can optionally contain a textual description of the album and of the disc, such as the album title, artist and publisher. This text can be stored in maximum 8 Language/Character set combinations.

The Area TOC's are located at the start and at the end of the corresponding Audio Area. The Area TOC's contain information on the Audio Area, such as the sampling frequency of the audio, the bit rate, the total playing time, the number of tracks, the track list, and optional text channels. The text channels may contain textual information such as track titles, performers, composer, etc.

File System

The HD layer of a Super Audio CD optionally contains a UDF and/or a ISO 9660 file system. The directory structure of a SACD disc is shown in figure 4. The files MASTER1.TOC, MASTER2.TOC and MASTER3.TOC are three identical copies of the MASTER TOC. The files 2C_AREA1.TOC and 2C_AREA 2.TOC correspond to two identical copies of the 2-channel stereo Area TOC, the files MC_AREA 1.TOC and MC_AREA2.TOC correspond to the Multichannel area TOC. The files TRACKnnn.2CH and TRACKppp.MCH correspond to the 2-Channel Stereo and Multichannel audio tracks respectively.

Audio Tracks

The audio tracks contain a multiplex of elementary streams. There are two types of elementary streams: the audio elementary stream, a sequence of audio frames, and the supplementary data elementary stream, a sequence of supplementary data frames. All frames represent a time period of 1/75 seconds. Each audio frame has a time code associated to it. Supplementary

Data Streams can contain text, graphics, still pictures etc. They can be perfectly synchronised with the audio, for e.g. Karaoke applications.

Audio Format

Basic Audio Format

The basic audio format for the Super Audio CD is the Direct Stream Digital (DSD) format, offering the ultimate in basic audio quality [3]. The DSD format transfers directly the upsampled bit stream to the disc, thereby eliminating the need for audio processing in the form of downsampling and upsampling filters and thus retaining the high timing accuracy required for an optimal stereo image. Although the superior quality of DSD is better demonstrated by listening than by showing figures, the basic specifications are given below.

A/D conversion process	1-bit Delta-Sigma
Sampling Frequency	64 times 44.1 kHz for 2-channel stereo 64 or 48 times 44.1 kHz for multichannel
Frequency range	to 100 kHz
Dynamic Range	120 dB (up to 20 kHz)

Two Channel Audio format

A stereo mix in the DSD format is contained in the 2-channel stereo area. The reproduction configuration is a conventional 2-channel set-up, with a left and a right loudspeaker or headphones. The contents of the 2-channel area is independent from the contents of the multichannel area, and may thus contain an optimised 2-channel mix, not required to be derived from the multichannel sound tracks. This ensures full artistic freedom in generating the mix.

Multichannel Audio format

A multichannel mix in the DSD format is contained in the multichannel area. The loudspeaker configuration is given in the area TOC. Currently, two, five and six channel loudspeaker configurations are defined. The possibility exists to add in the future new loudspeaker configurations, if acoustic research gives rise hereto.
In figure 5, one of these loudspeaker configurations is shown. The channels are Left, Centre, Right, Left Surround, Right Surround.

Lossless Coding

Lossless coding can be applied to DSD audio frames to achieve a more efficient storage [4]. Lossless coding is a method whereby the process of encoding and decoding is completely transparent, i.e. a bit exact replica of the original signal is obtained at the decoder side. The process is depicted in figures 6 and 7. Typical gains that can be achieved on a DSD stream are in excess of a factor two. This technique allows storage of 74 minutes of DSD 2-channel stereo and 74 minutes of DSD multichannel on one HD layer.
The application of lossless coding can be decided upon on a frame by frame basis. The necessary signalling to the decoder is contained in the header of the audio frames. If a disc does not contain lossless coded frames at all, this is signalled on the disc by indicating the 'plain DSD' profile in the area TOC.

Buffering

As stated above, the audio tracks contain a multiplex of elementary streams. The multiplexed stream has a constant bit rate to avoid the noise of speed changes or jump-backs in the player. To optimally fit the variable sized lossless coded frames into a fixed bit rate, a buffer model is used, smoothening the bit rate at places where few frames need more bits and introducing stuffing at places where the lossless coding generates smaller frames. Every sector starts with a header containing information on the type of data present in the sector. The administration of the buffer model is also contained in these headers.

Extra Data Area

The Extra Data Area contains information that can only be retrieved via the file system. The format of the data in the Extra Data Area is not specified in the Super Audio CD specification, to keep maximum flexibility.

References

1. ten Kate, R., "Disc-technology for super Quality audio applications", 103rd AES Convention, September 1997, New York
2. Philips and Sony Corporation, "System Description Compact Disc Digital Audio (Red Book)"
3. Nishio, A., et al., "Direct Stream Digital Audio System", 100th AES Convention, May 1996, Copenhagen
4. Fons Bruekers, Werner Oomen, René van der Vleuten, Leon van de Kerkhof, "Improved Lossless Coding of 1-bit Audio Signals", 103rd AES Convention, September 1997, New York

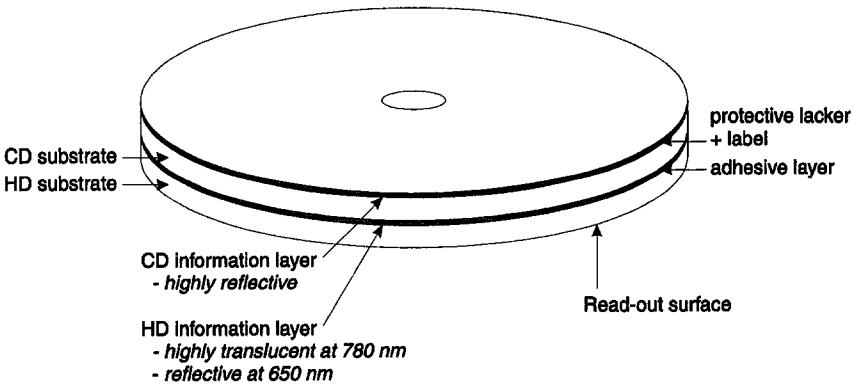

Figure 1: the hybrid disc

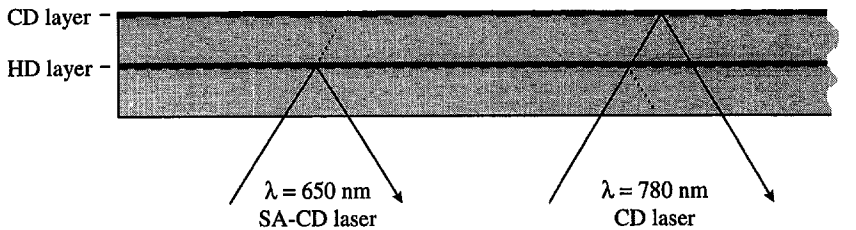

Figure 2: CD compatibility

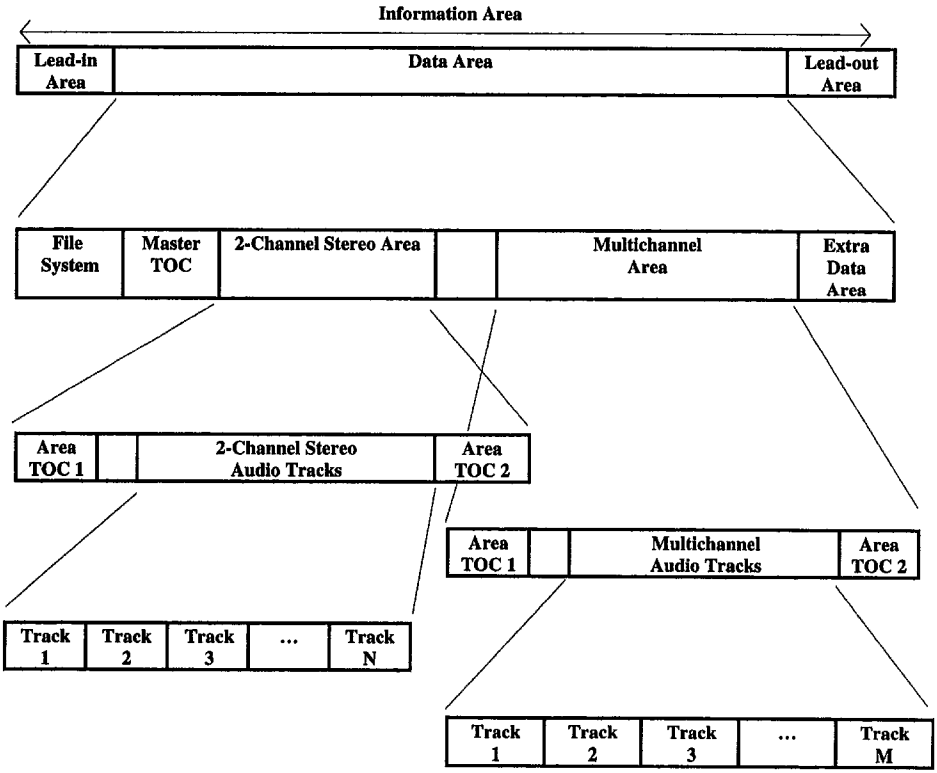

Figure 3: The structure of the HD layer

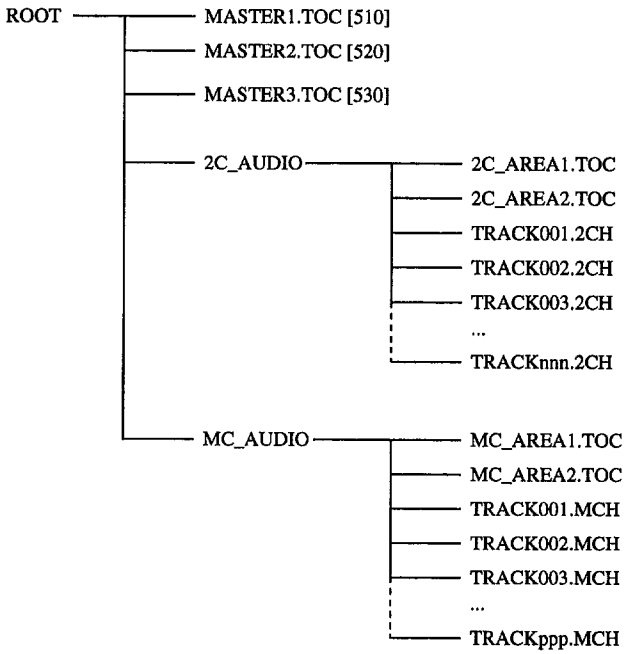

Figure 4 : The File System

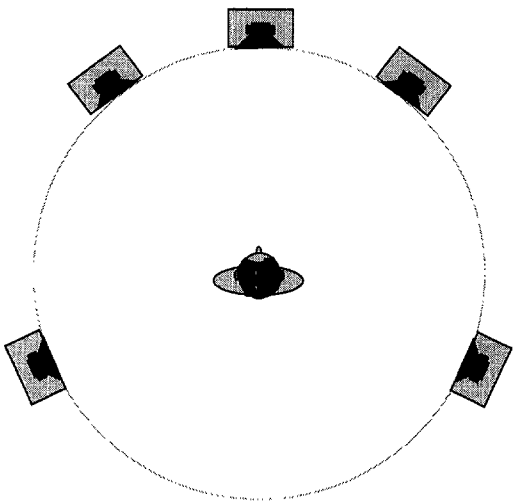

Figure 5: One of the specified loudspeaker set-ups

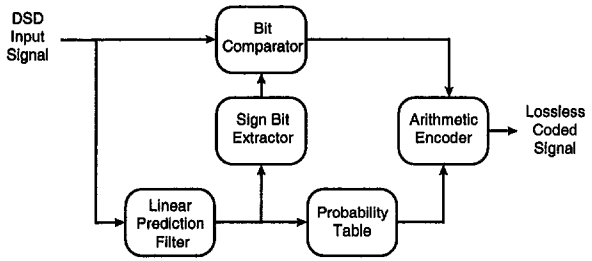

Figure 6: the Lossless Encoder (LLC)

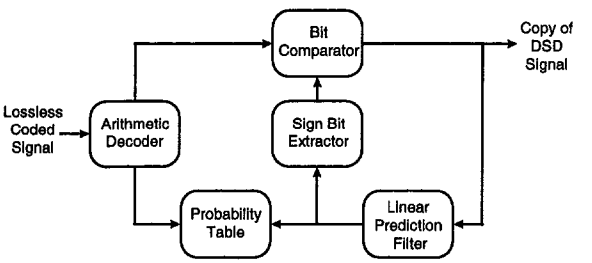

Figure 7: the Lossless Decoder (LLC^{-1})