DVD: Application of Hierarchically Encoded Surround Sound – including Ambisonics

Acoustic Renaissance for Audio

Confidentiality: Public Domain. Version 0.9a

This Proposal is addressed to members of the DVD Consortium [1]*, to WG-4 of the DVD Consortium, to members of IFPI, RIAA, RIAJ and of the Audio Engineering Society.*

 jrs@meridian-audio.com

0. Summary

This document supplements other documents from the *ARA* concerning audio for both the Video DVD (DVDV) and a pure-audio DVD (DVDA). [1]

We suggest that the standards for DVD Video [6], and any new standard for DVD Audio incorporate a flag to indicate multichannel material for which the original was in *Ambisonic* format, and a second flag that indicates when the multichannel information is hierarchically encoded – and therefore – although the stream carries suitable loudspeaker feeds, more surround information can be extracted by a suitable decoder.

Adding indicators for hierarchical or *Ambisonic* material does not in any way limit the use of the DVD. Producers are free to make multichannel recordings in any way they wish.

We believe it should be considered urgently.

1. Actions Required

We are asking for 2 additional flags in the DVD stream. These indicate the origin of the data in the stream. Details are given in Section 5.

2. Extent of discussion and authorship

This document is a proposal from members of *Acoustic Renaissance for Audio,* a free body dedicated to advancing audio quality. The signatories, their advisors and all affiliations are appended in section 7.

[1] The DVD standard is published by, and copyright of: *Hitachi Ltd., Matsushita Electric Industrial Co. Ltd., Mitsubishi Electric Corporation, Philips Electronics N.V., Pioneer Electronic Corporation, Sony Corporation, Thompson Multimedia, Time Warner Inc., Toshiba Corporation, Victor Company of Japan Inc.*

3. Ambisonics coding

Ambisonics is a mature and comparatively complete technology and should be allowed for in multichannel audio carriers.

Compared with other methods of coding, *Ambisonics* has these benefits:

- The surround sound field is described independent of the target reproducing loudspeaker layout.
- The method allows realistic re-creation of natural sound fields.
- Decoders can be constructed for any number and position of loudspeakers.
- Compatible with existing systems

Ambisonics represents a sound-field in what is described as 'B-format'. Normal *Ambisonics* B-format contains 4 independent channels of information, which give a 3-axis representation of the sound at the recording position. The signals are namely:

W	Mono sound pressure
Y	Left–Right velocity component
X	Front-Back velocity component
Z	Up–Down velocity component (optional)

An *Ambisonic* decoder processes these signals to provide any number of loudspeaker feeds.

For example, the three signals WXY are quite sufficient to provide full horizontal surround sound when processed to five (or more) loudspeakers. As such, this format is very efficient in data rate and in flexibility in decoding.

If more channels are available, then additional signals have been suggested to improve angular resolution.

An option to supply *Ambisonics* in DVD is to store *Ambisonic* material on the disc as loudspeaker feeds. In this method, B-Format is decoded at the mastering stage using a standard decoder layout to provide loudspeaker feeds for 5 speakers (or 6 speakers, one above the listener).

This multichannel presentation should then be flagged as '*Ambisonic origin*'. Sophisticated decoders can read the flag, reconstitute B-format from the standard decoder parameters and then re-decode exactly for the target system.

4. Multichannel Coding

The B-Format signals can also be represented by the following signals:

M	Mono sound pressure
S	Left–Right difference component
B	Front-Back difference component
V	Up–Down difference component (optional)

MSBV can also be derived from *any multichannel material* (not necessarily *Ambisonic*). Gerzon has also shown that there is an extremely flexible means of transcoding such material between different numbers of loudspeakers.

More importantly, Gerzon has shown [4] that either the WXYZ or MSBV formats can be transformed to a hierarchy of speaker feeds, from which the original signals can be recovered. This transform then results in combinations of the following multi-speaker signal sets:

L3 Compatible Left speaker feed

R3 Compatible Right speaker feed

C3 Compatible Centre speaker feed, *or*

B3 Compatible Back speaker feed

V3 Compatible Top speaker feed

For example, the set L3 + R3 + C3, allows a single audio stream on the DVD to convey, without processing, a suitable result using the three front loudspeakers. A more sophisticated decoder can recover the full three-dimensional sound field in an efficient manner from these feeds.

This hierarchical scheme can be extended to higher orders as described in [4].

This second option allows multichannel playback without an *Ambisonic* decoder, but has a higher data-rate as more channels are conveyed (5 instead of 3 for horizontal *Ambisonics*; 6 instead of 4 for with-height *Ambisonics* i.e *Periphony*).

5. Representing Hierarchies and Ambisonics on DVD

We suggest that two descriptors are needed to allow correct use of Ambisonics.

5.1 Hierarchical Loudspeaker feeds

Changes to DVD Spec. 1.0:

1. Channel assignment mode in Application Information:

- Add Multichannel Type 2: Hierarchical Format recording (See E.1.2)
- Suggestion: 010b: Multichannel Type 2: Hierarchical Format

For which the Audio channel contents descriptor needs to be like E.1.2.1 and as follows:

Code	Number Channels	1	2	3	4	5	6	7	8
	3	L3	R3	C3					
	3	L3	R3	B3					
	4	L3	R3	C3	V3				
	4	L3	R3	B3	V3				
	5	L5	R5	C5	SL5	SR5			
	6	L5	R5	C5	SL5	SR5	V5		

This mode indicates to a sophisticated decoder that more exact surround information is available by suitable processing, but that the signals may be played directly to loudspeakers.

5.2 Ambisonics as Loudspeaker feeds

Changes to DVD Spec. 1.0:

1. Channel assignment mode in Application Information:

- Add Multichannel Type 3: Multichannel from Ambisonic material
- Suggestion: 011b: Multichannel Type 3: Ambisonic Speaker Feeds

This mode is identical to Multichannel Type 1, except that it is indicated that the speaker feeds originate from an Ambisonic source and that therefore sophisticated replay devices can recover the B-Format. If agreed, we would suggest standard speaker layouts.

6. Rights

Ambisonics is a trade-mark of *Nimbus Recording*. There may or not be license issues for building decoders or encoders using the Hierarchical structure. We are not aware that the flag carries any limitation.

7. Authorship

Bob Stuart. Chairman and Technical Director, Meridian Audio Ltd. Visiting Fellow Essex University, Fellow *Audio Engineering Society*, Member *Acoustical Society of America*, Chairman *Acoustic Renaissance for Audio*, Member *XtraBits*, Member *Technical Subcommittee National Sound Archive*, Member *IEE* and *IEEE*. Member *ADA* committee of *Japan Audio Society*.

Peter Craven. Consultant. Member *Audio Engineering Society*, Member *XtraBits*.

Takeo Yamomoto. Chief Auditor, Pioneer Electronic Inc., Member *Audio Engineering Society*, Founder *Japan Acoustic Renaissance for Audio*, Member *ADA* committee of *Japan Audio Society*.

Professor Malcolm Hawksford. University of Essex. Fellow *Audio Engineering Society*, Fellow *Institute of Acoustics*, Fellow *IEE*., Member *Acoustic Renaissance for Audio*.

Chris Travis. Sonopsis Ltd. Member *Audio Engineering Society*.

Tony Griffiths. Consultant. Fellow *Audio Engineering Society*, Member *Acoustic Renaissance for Audio*, Chairman *Technical Subcommittee National Sound Archive*, Member *IEE*, Member *Royal Television Society*.

Michael Gerzon. (Deceased). Gold Medallist and Fellow *Audio Engineering Society*, Former Member *XtraBits*, Former Member *Acoustic Renaissance for Audio*.

8. Bibliography and references

1 *Acoustic Renaissance for Audio*, 'A Proposal for High-Quality Application of High-Density CD Carriers', private publication (April 1995).

2 Gerzon, M.A. 'Periphony: With-Height Sound Reproduction' *J. Audio Eng. Soc*., **21**, 2–10 (Jan/Feb 1973)

3 Gerzon, M.A. 'Ambisonics in Multichannel Broadcasting and Video' *J. Audio Eng. Soc*., **33**, 859–871 (Nov 1985)

4 Gerzon, M.A. 'Hierarchical System of Surround Sound Transmission for HDTV', *92nd AES Convention*, Vienna, preprint 3339 (March 1992)

5 Gerzon, M.A. and Barton, G.J. 'Ambisonic Decoders for HDTV', *92nd AES Convention*, Vienna, preprint 3345 (March 1992)

6 Toshiba et al., 'DVD Specifications for Read-Only Disc', Version 0.9 (April 1996)

9. Contact Addresses

Acoustic Renaissance for Audio, Meridian Audio, Stonehill, Stukeley Meadows, HUNTINGDON PE18 6ED, England. **Phone**: +44 1480 52144 **Fax**: +44 1480 451587

Email: ara@meridian-audio.com **Web:** http://www.meridian-audio.com/ara

10. Trademarks, etc.