

Matrixed Surround sound in an MPEG digital world

D. J. Meares

*BBC Research and Development Department, Kingswood Warren, Tadworth,
Surrey KT20 6NP, U. K.*

G. Theile

Institut für Rundfunktechnik, Floriansmühlstrasse 60, D-80939 München, Germany

Via such systems as the BBC's NICAM and the German dual-carrier FM sound systems, listeners to television have been able to receive Dolby Stereo, a surround sound service common to cinema and home theatre audiences. Evidence exists to show that an increasing number of television listeners are now equipping themselves with the appropriate decoding equipment, e.g. a Pro-Logic decoder, to enable them to listen in surround sound. Thus any move by television companies to Digital Terrestrial (or non-terrestrial) Television broadcasting which uses bit-rate-reduction technologies, must consider the wishes of this section of their audience.

This paper, written on behalf of the Audio Working Group of EU1187 PGII, considers the options for broadcasting Dolby Stereo via MPEG coding as an initial route to providing surround sound services.

Although this paper, for convenience, is written specifically in the context of the Dolby Stereo system, much of what is stated is equally applicable to other two-channel based sound systems, such as Sensora Sound, Ambisonics, Circle Surround, etc.

0. INTRODUCTION

For approximately 20 years Dolby have been providing the film industry in particular with a form of surround sound which they called Dolby Stereo [1]. Over a period of about ten years, the corresponding consumer version "Dolby Surround¹" has become a world leader and was in fact the only commercial success² from the 1970's experiments into quadraphony. To provide the surround sound the incoming signals are encoded through a phase/amplitude matrix to create stereo-like signals Lt and Rt. These can be treated as stereo signals for the purpose of listening, recording or transmission. When it is desired to recreate the surround sound signals they have to be decoded and these days the normal decoder

is what is now known as a Pro-Logic decoder. This arrangement is shown in Figure 1.



Figure 1 Conventional handling of Dolby Stereo

Though the system was good for its time, it has to be acknowledged that it brought with it constraints on the way in which the surround sound dimension could be used. These constraints are normally minimised by monitoring, at the mixing stage, through the coder and decoder, but they are the reason why the film and broadcasting industries have been seeking alternative ways of providing surround sound in the digital age.

Despite the newer proposals, such as DTS, MPEG-2, AC-3, MPEG AAC etc. [2, 3, 4, 5], broadcasters and others still need to be able to handle Dolby Stereo, if only because of their archive programmes. For digital

¹ For convenience the term Dolby Stereo will be used for the rest of this article to differentiate the 2-channel matrix system from the 5-channel AC3 digital system.

² It should be noted that the Ambisonics system had similar origins, but is certainly not so well known nor in such widespread use.

broadcast applications there are several options and it is the purpose of this paper to describe those options and to recommend the more cost effective for further consideration.

In the first instance, the situation that now exists for stereo networks using MPEG-1 will be described. Next will be described the means for handling archive Dolby Stereo material within the MPEG-2 world. Finally, options will be discussed where specific Dolby Stereo decoding can be facilitated at the same time as discrete multichannel services.

1. STEREO MPEG-1 ENVIRONMENTS

For those users of stereo digital compression systems, such as MPEG-1 [6], one might assume that the Dolby Stereo signals would pass unchanged through stereo bit rate reduction coding, at least no more changed than any other stereo signals. However, one must consider that after MPEG-1 decoding the signals should be decodeable by the Pro-Logic decoder (see Figure 2). If, additionally, one remembers that the decoder is looking for certain phase and amplitude signatures to enable it to steer the sound to the correct spatial location, one can then consider whether the coding options for MPEG-1 enable the phase and amplitude relationships to be preserved.

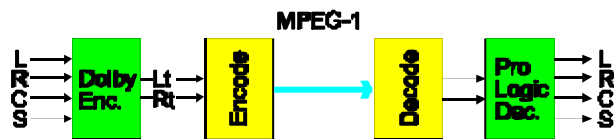


Figure 2 Dolby Stereo via an MPEG-1 channel

In earlier studies of the problem, members of the Eureka 1187 (ADTT) project Audio Subgroup determined that, with the exception of one MPEG coding option, it is possible to maintain the Dolby coding signatures [7, 8]. Joint Stereo coding is the one feature of MPEG, and other bit rate reduction systems, that causes problems. The principle of this coding option is to save bit rate by exploiting the ear's insensitivity to phase information at frequencies greater than 2 kHz and by coding only the amplitude differences between the channels at those frequencies. On the other hand, the phase information is important to the Dolby system at frequencies below 7 kHz, because it is needed for dematrixing the 7 kHz lowpass filtered surround signal, S. The phase information below 7 kHz must, therefore, be preserved through the MPEG coding and decoding. As a result, Dolby Stereo capable MPEG encoders do not employ the Joint Stereo coding technique at frequencies below 7 kHz. In cases where the MPEG encoder is not Dolby Stereo capable, the Joint Stereo mode should not be used.

Considering this, the observation which must be made is that if the signals are to be decoded through Pro-Logic and distributed into four channels, there is some evidence that a higher stereo bit rate is necessary for a given quality than if only stereo were being considered. This however is part of the much greater and much more difficult debate on what constitutes an adequate quality service. It is not the subject of this paper.

2. THE NEW ERA OF MULTICHANNEL BIT RATE REDUCTION CODING

Multichannel digital coding, e.g. MPEG-2, brings with it new options for providing surround sound services to the listening public. By handling all channels much more independently, the constraints of the Dolby Stereo system can be eliminated. Figure 3 shows the digital option.

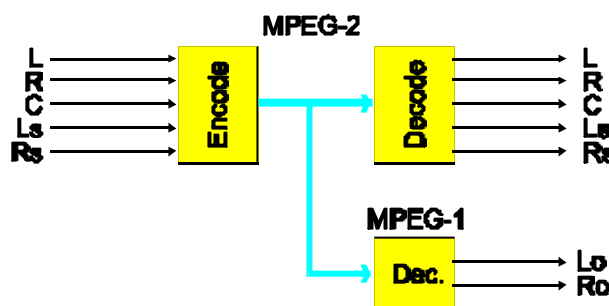


Figure 3 MPEG-2 discrete digital surround sound

All proposals [see for example 9, 10, 11] for surround sound start with the acknowledgement that five source and destination channels are required³. As shown above, these are Left, Right, Centre, Left surround and Right surround channels. Tests have shown that this arrangement is far superior to the four channel option [12]. In the case of MPEG-2, these source signals are encoded into a single bitstream prior to transmission. Because of the backwards compatibility of MPEG-2 to MPEG-1 [13], the bitstream can be decoded correctly by both MPEG-1 and MPEG-2 decoders.

For this paper, the compatibility of MPEG-2 with Dolby Stereo Lt/Rt material has to be considered. The basic question is: "How can we transmit Lt/Rt material in a multichannel MPEG-2 broadcasting chain?"

³ The options for a low frequency effects (LFE) channel and additional services have been omitted from this paper as they are not germane to the arguments presented here.

2.1. Dolby Stereo decoding prior to MPEG-2 encoding

If we already have a multichannel coding system, one has to consider whether it is appropriate to decode all sources to a single format and then code them for transmission in a standardised way. This will, after all, ensure uniformity at the point of reception and the requirement for only a single form of decoder in the consumers' equipment. It also means that any improved method of decoding Dolby Stereo could be provided to all consumers by changing only the hardware located at the transmission point. This option is illustrated in Figure 4.

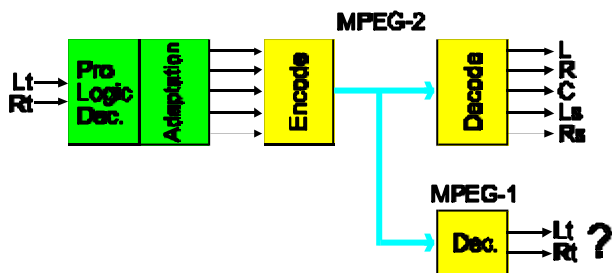


Figure 4 Decoding Dolby Stereo prior to compression

The arrangement would be that, at the point of origin, any Dolby Stereo signals would be Pro-Logic decoded. The resultant signals would then be modified slightly to resolve the conflict between the mono surround provided by the Dolby system and the stereo surround required for transmission. This could be as simple as replicating the signal and attenuating both surround feeds by 3 dB to keep the overall surround power correct. The decoded signals would then be treated as any 5 channel source, providing surround sound through the MPEG-2 decoder in the normal way.

What though of the compatible MPEG-1 reception? By design, the MPEG-1 receiver is supposed to be provided with data that will generate a stereo service. However, in this scenario the receiver is being provided with signals that are a mix down of the Dolby decoded signals, which were themselves generated by a variable ProLogic decoder matrix designed to steer the signals in a surround sound presentation. Gain-riding is therefore taking place, and delay is introduced into the surround channel, both measures being designed only to be subjectively balanced in surround. In limited tests it has been found that the mixed down Dolby decoded signals present a very different sound from that created by the initial Dolby encoded signals. Changes to balance, the sense of spaciousness and image location are clearly evident. If the programme producer intended the listener to hear the sound as generated by the encoded signals, it is inappropriate to make radical changes to the subjective effect without very good reason.

Two other points also rule out this method of handling Dolby Stereo. Firstly, the MPEG-1 receiver is not going to receive signals that can be decoded by a Pro-Logic decoder. This means that many listeners who have set up a Home Theatre based on Dolby Pro-Logic will lose that service unless they go out and buy an MPEG-2 receiver. This may or may not be acceptable to the service provider.

Much more importantly, the demand for bit rate on cable, satellite or terrestrial broadcasting is going to be such that all options for conserving that bit rate must be considered. If we are starting off with what is basically a stereo form of signal, why decode it to five channels before bit rate reduction coding? All evidence to date shows that five channel coding requires more bit rate than two channel coding, even without using the Joint Stereo coding mode.

An additional option could be to transmit, in MPEG-2, a combination of Lt and Rt as the backwards compatible components together with Co, Ls and Rs from the Prologic decoder to provide the 5-channel service. This would indeed provide the MPEG-1 listener with signals suitable for subsequent Dolby Stereo decoding, but would not be capable of providing directly a 5-channel service without specific dematrixing. This then would not be a service compatible with a standard MPEG-2 receiver, the 5-channel derived signals would be subject to gain-riding artefacts and it would still require the bit rate of a 5-channel system rather than that of a stereo system.

It is thus recommended that the scheme shown in Figure 4 should not be adopted.

2.2. Dolby Stereo decoding in combination with MPEG-2 (and MPEG-1) decoding

Figure 5 shows the arrangement by which a programme source originating in Dolby Stereo format should be transmitted or otherwise conveyed to the listener. The Lt/Rt signals are encoded for transmission via a standard MPEG-2 coder⁴ working in its 2-channel mode.

⁴ The assumption here is that the service provider is offering a multichannel service into which he/she wishes to insert a Dolby Stereo programme.

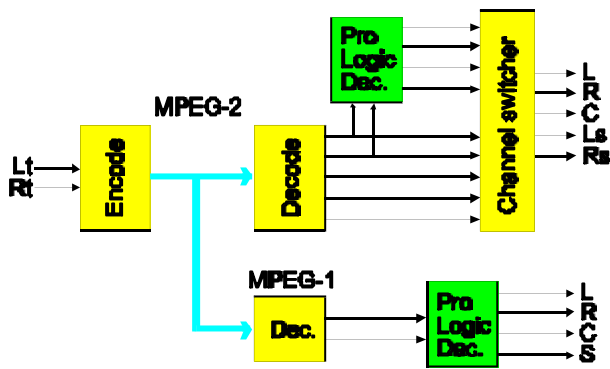


Figure 5 Preferred handling of Dolby Stereo with MPEG-2 coding

The MPEG-2 coding considerations specific to the conveyance of Dolby Stereo signals are described in [6, 7]. The MPEG-1 decoder regenerates the stereo signals and can then be coupled to a Pro-Logic decoder for surround listening. The MPEG-2 decoder will receive a 'flag' from the encoder, telling it that it is receiving a Dolby Stereo signal and can be configured to apply the Lt/Rt signals to a built-in Pro-Logic decoder. When 5-channel source material is being encoded for transmission, the Dolby Stereo flag is removed and the MPEG-2 decoder is configured to provide direct feeds of the decoded signals, bypassing the Pro-Logic decoder.

At first sight the provision of both MPEG-2 and Pro-Logic decoding in the receiver might seem excessive and costly. However, it should be remembered that for the foreseeable future consumers are going to want to continue playing their VHS tapes and other programme sources. Where these are Dolby Stereo encoded, the consumer can legitimately expect to be able to decode them at home to provide Dolby surround.

Additionally, the Pro-Logic decoder is currently available on a single chip and does not add significantly to the manufacturer's costs. Further in the future, the MPEG decoding may be achieved within DSP chip sets of sufficient power to be reconfigurable to provide Pro-Logic decoding as an alternative⁵. Thus, though Figure 5 depicts the Pro-Logic decode function as a separate entity, it could in the course of time be just another software component of the overall receiver/decoder.

The benefit, throughout all this, is that during a Dolby Stereo programme the bit rate required of the transmission chain or the recording medium, in such things as DVD or CD-I, is that of a stereo service rather than that of a discrete 5-channel service.

3. COMBINED MPEG AND DOLBY STEREO PROVISIONS

Up to this point in the paper, 5channel MPEG and Dolby Stereo provisions have been considered as separately timed options. However if one is transmitting a 5-channel programme via MPEG-2, it is possible simultaneously to provide a Dolby Stereo service via MPEG-1. There is a provision within the MPEG-2 Audio specification to code the signals such that the backwards compatible components of the bitstream can be decoded via Pro-Logic, to provide Dolby surround, whilst the full bitstream is decodable via MPEG-2 decoding, to provide the full surround sound service. This is depicted in Figure 6.

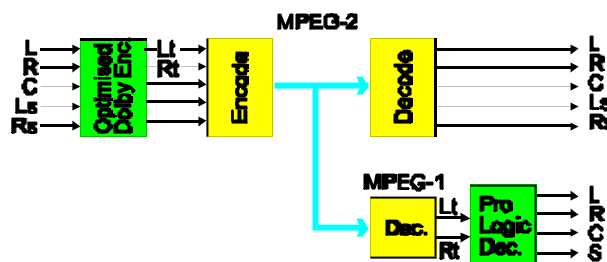


Figure 6 Combined Dolby and MPEG-2 coding

The special provisions within MPEG-2 coding eliminate some of the limitations of conventional analogue Dolby coding that were a requirement of optical film sound tracks. These details include bandlimiting and the noise companding of the surround channels, both of which are avoided in this scenario. The end product of this will be better Dolby Stereo presentations for both stereo and surround reproduction.

It should be noted, however, that the constraints of Dolby Stereo are such that specially optimised mixes of the programmes are normally required at the point of origination to overcome the limitations of Dolby Stereo. If one is starting with an already balanced 5-channel mix, this will not necessarily be optimum for Dolby Stereo⁶. The end product will be different sound presentations from the Dolby Stereo and the full 5-channel services, but at least both sets of listeners will receive some form of surround sound.

The original 5channel presentation is, however, not changed at all by the above process and the decoded representation of the original sound sources is precisely what was intended by the producer.

⁵ Interestingly, this is already the case for at least one domestic Dolby AC-3 decoder with Pro-Logic option.

⁶ This is a generalised problem for Dolby Stereo, which is equally applicable to the circumstances of MPEG coding, AC-3 coding etc..

4. CONCLUSIONS

It has been shown that the introduction of digital multichannel compressed sound services need not result in the loss of existing Dolby Stereo services. The inclusion of Dolby Stereo options requires the use of Dolby Stereo capable MPEG encoders. This takes a slightly higher bit rate (compared to just two channel stereo) for the same audio quality. It has been recommended that the Dolby Stereo signals should not be decoded prior to transmission, but should be MPEG coded and transmitted as stereo signals. Provisions within the receiver have been described and an option for combined services based on 5-channel source material has been outlined.

5. ACKNOWLEDGEMENTS

The authors would like to thank the management of the BBC and IRT for permission to publish this paper.

6. REFERENCES

- 1 Hull, J., 1994. Surround sound: past, present and future. Dolby Laboratories Inc. Brochure. 1994.
- 2 Fox, B., 1995. Technology. Hi-Fi News and Record Review. February 1996. p17.
- 3 ISO/IEC, 1994. Generic coding of moving pictures and associated audio: Audio Part. International Organisation For Standardisation, Coding Of Moving Pictures and Audio. ISO/IEC IS 13818-3. November 1994.
- 4 Forshay, S., 1993. Dolby AC-3 Multichannel Digital Audio Compression System, Encoder Algorithm Description. ISO/IEC MPEG document number 93/972.
- 5 ISO/IEC, 1997. Generic coding of moving pictures and associated audio: Advanced Audio Coding (AAC). International Organisation For Standardisation, Coding Of Moving Pictures and Audio. ISO/IEC IS 13818-7. April 1997.
- 6 ISO, 1992. Coding of Moving pictures and associated audio for digital storage media at up to 1.5 Mbit/s: Audio Part (11172-3). Draft International Standard 11172. 1992.
- 7 Nielsen, S. H., Stoll, G., v.d. Kerkhof, L.; 1994. Perceptual Coding of Matrixed Audio Signals. Proceedings of the 96th AES Convention, Amsterdam, February 1994, Pre-print 3867
- 8 Nielsen, S. H., Stoll, G., 1994. ISO/MPEG Layer II coding of Dolby surround signals. Proceedings of 18th Tonmeistertagung, Karlsruhe, Germany. November 1994.
- 9 ITU-R. 1993. Multichannel stereophonic sound system with and without accompanying picture. ITU-R Recommendation 775 (Rev1), Doc 10/63 (November, 1993).
- 10 SMPTE. 1991. Proposed recommended practice: loudspeaker placements for audio monitoring in high definition electronic production. Document N15.04/152-300. SMPTE, New York December 1991
- 11 Meares, D. J., Theile, G., 1992. High quality sound for high definition sound. EBU book 'Advanced techniques for satellite broadcasting of digital HDTV at frequencies around 20 Ghz.' ITU WARC92 Conference, Torremolinos, February 1992. pp 65 to 78.
- 12 Meares, D. J., 1991. High quality sound for high definition television. Proceedings of the AES 10th International Conference 'Images of audio', Kensington, London, 7 - 9 September, 1991. pp 163 to 177
- 13 ten Kate, W. Compatibility Matrixing of Multichannel Bit-Rate-Reduced Audio Signals. J. Audio Eng. Soc., Vol. 44, no. 12, 1996, pp. 1104 to 1119