



ATSC

ADVANCED TELEVISION
SYSTEMS COMMITTEE

ATSC Digital Television Standard, Part 6 – Enhanced AC-3 Audio System Characteristics

Doc. A/53 Part 6:2013
7 August 2013

Advanced Television Systems Committee
1776 K Street, N.W.
Washington, D.C. 20006
202-872-9160

The Advanced Television Systems Committee, Inc., is an international, non-profit organization developing voluntary standards for digital television. The ATSC member organizations represent the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

Specifically, ATSC is working to coordinate television standards among different communications media focusing on digital television, interactive systems, and broadband multimedia communications. ATSC is also developing digital television implementation strategies and presenting educational seminars on the ATSC standards.

ATSC was formed in 1982 by the member organizations of the Joint Committee on InterSociety Coordination (JCIC): the Electronic Industries Association (EIA), the Institute of Electrical and Electronic Engineers (IEEE), the National Association of Broadcasters (NAB), the National Cable Telecommunications Association (NCTA), and the Society of Motion Picture and Television Engineers (SMPTE). Currently, there are approximately 150 members representing the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

ATSC Digital TV Standards include digital high definition television (HDTV), standard definition television (SDTV), data broadcasting, multichannel surround-sound audio, and satellite direct-to-home broadcasting.

Note: The user's attention is called to the possibility that compliance with this standard may require use of an invention covered by patent rights. By publication of this standard, no position is taken with respect to the validity of this claim or of any patent rights in connection therewith. One or more patent holders have, however, filed a statement regarding the terms on which such patent holder(s) may be willing to grant a license under these rights to individuals or entities desiring to obtain such a license. Details may be obtained from the ATSC Secretary and the patent holder.

Revision History

A/53 Part 6:2007	3 January 2007	Initial version of A/53 Part 6 approved. Previously Part 6 was Annex G of A/53.
A/53 Part 6:2010	6 July 2010	First revision of A/53 Part 6 approved.
A/53 Part 6:2013	7 August 2013	Second revision of A/53 Part 6 approved.

Table of Contents

1. SCOPE	6
2. REFERENCES	6
2.1 Normative References	6
2.2 Informative References	6
3. COMPLIANCE NOTATION	7
3.1 Treatment of Syntactic Elements	7
3.2 Symbols, Abbreviations, and Mathematical Operators	7
4. SYSTEM OVERVIEW	7
5. CONSTRAINTS ON E-AC-3 CODING	8
5.1 Syntactical Element Constraints	8
5.2 Sampling Rate	8
5.3 Audio Coding Modes	9
5.4 Constraints for Audio Mixing Applications	9
5.5 Dialog Level	10
5.6 Dynamic Range Compression	10
5.6.1 Dynamic Range Compression – “Line Mode”	10
5.6.2 Dynamic Range Compression – “RF Mode”	10
6. INTRODUCTION TO E-AC-3 AUDIO SERVICES	10
6.1 Options for Providing Video Description	10
7. CONSTRAINTS ASSOCIATED WITH SPECIFIC AUDIO SERVICE TYPES	11
7.1 For service_type Values of 2 and 3	11
7.1.1 Complete Main (CM) Audio Service	11
7.1.2 Main Audio Service, Music and Effects (ME)	11
7.1.3 Visually Impaired (VI)	11
7.1.4 Hearing Impaired (HI)	11
7.1.5 Additional bitstream Constraints	11
7.2 For service_type Value 7	11
7.2.1 Complete Main (CM) Audio Service	12
7.2.2 Main Audio Service, Music and Effects (ME)	12
7.2.3 Visually Impaired (VI)	12
7.2.4 Hearing Impaired (HI)	12
7.2.5 Dialog (D)	12
7.2.6 Commentary (C)	12
7.2.7 Voice-Over (VO)	12
7.2.8 Additional Bitstream Constraints	13
7.2.9 Stream Info Details for stream_type 0x87, service_type value 7	14
8. MIXING OF MAIN AND ASSOCIATED AUDIO SERVICES	14
8.1 Introduction	14
8.2 Mixing Metadata	14
8.2.1 Control Smoothing	15
8.2.2 Output Limiting	15
8.3 Examples of Decoding and Mixing Configurations	15
8.3.1 Mixing of AC-3 Main with E-AC-3 Associated Audio	15
8.3.2 Mixing of main and associated services in separate bitstreams	17

8.3.3	Mixing of 5.1-channel main with 3-channel associated audio	20
9.	CONVERSION OF E-AC-3 INDEPENDENT SUBSTREAM 0 TO AC-3	22
10.	ENCODING AND DECODING OF 7.1-CHANNEL PROGRAMS	24
10.1	Introduction	24
10.2	Encoding Process	24
10.3	Bitstream Structure	25
10.4	Decoding Processes	26
10.4.1	7.1 Decode Process	26
10.4.2	5.1/2.0 Decode Process	27

Index of Tables and Figures

Table 7.1 stream_info_details() Syntax for Stream Type 0x87 (E AC-3)	14
Table 8.1 AC-3_audio_stream_descriptor Configuration – Main Audio Service	17
Table 8.2 E-AC-3 audio stream descriptor Configuration – Associated VI Audio Service	17
Table 8.3 E-AC-3 audio stream descriptor Configuration – Main Audio Service	19
Table 8.4 E-AC-3 audio stream descriptor configuration – associated VO audio service	20
Table 8.5 E-AC-3_audio_stream_descriptor configuration – ME and associated D audio services	22
Figure 4.1 Audio subsystem in the digital television system.	8
Figure 8.1 Mixing an AC-3 main with an E-AC-3 associated service.	16
Figure 8.2 Mixing of a downmixed multichannel main with mono associated audio.	18
Figure 8.3 Mixing a 5.1-channel Main with a 3-channel Associated.	21
Figure 9.1 E-AC-3 to AC-3 conversion process.	23
Figure 10.1 Example 7.1-channel program configuration.	24
Figure 10.2 7.1 channel E-AC-3 encoder.	25
Figure 10.3 Fragment of an E-AC-3 stream structure comprising a 7.1-channel Main Audio Service.	26
Figure 10.4 Fragment of an E-AC-3 stream structure comprising a 7.1-channel Main Audio Service and a 2-channel Associated Audio Service.	26
Figure 10.5 7.1-channel E-AC-3 decoder.	27
Figure 10.6 5.1-channel E-AC-3 decoder.	28

A/53 Part 6:2013 – Enhanced AC-3 Audio System Characteristics

1. SCOPE

This Part of A/53 establishes a set of constraints on A/52 [1] for, and normative provisions for, the coding of audio as Enhanced AC-3 (E-AC-3).

2. REFERENCES

All referenced documents are subject to revision. Users of this Standard are cautioned that newer editions might or might not be compatible.

2.1 Normative References

The following documents, in whole or in part, as referenced in this document, contain specific provisions that are to be followed strictly in order to implement a provision of this Standard.

- [1] ATSC: “Digital Audio Compression (AC-3, E-AC-3) Standard,” Doc. A/52:2012, Advanced Television Systems Committee, Washington, D.C., 17 December 2012.
- [2] ANSI: “Specification for Sound Level Meters,” Doc. ANSI S1.4-1983 (R 2001) with Amd.S1.4A-1995, American National Standards Institute, Washington, D.C., 2001.
- [3] ITU: “Algorithms to measure audio programme loudness and true-peak audio level,” Recommendation ITU-R BS.1770-3, International Telecommunications Union, Geneva, 2012.

2.2 Informative References

The following documents contain information that may be helpful in applying this Standard.

- [4] AES: “AES Standard for digital audio engineering—Serial transmission format for two-channel linearly represented digital audio data,” Doc. AES3-2009, Audio Engineering Society, New York, N.Y., 2009.
- [5] ATSC: “ATSC Digital Television Standard, Part 3 – Service Multiplex and Transport Subsystem Characteristics,” Doc. A/53 Part 3:2013, Advanced Television Systems Committee, Washington, D.C., 7 August 2013.
- [6] ATSC: “ATSC Digital Television Standard, Part 5 – AC-3 Audio System Characteristics,” Doc. A/53, Part 5:2010, Advanced Television Systems Committee, Washington, D.C., 6 July 2010.
- [7] ATSC: “Digital Television Standard, Part 1 – Digital Television System,” Doc. A/53 Part 1:2013, Advanced Television Systems Committee, Washington, D.C., 7 August 2013.
- [8] ATSC: “Recommended Practice – Techniques for Establishing and Maintaining Audio Loudness for Digital Television,” Doc. A/85:2013, Advanced Television Systems Committee, Washington, D.C., 12 March 2013.
- [9] SMPTE: SMPTE ST 428-3:2006 D-Cinema Distribution Master – Audio Channel Mapping and Channel Labeling
- [10] ITU: “Multichannel stereophonic sound system with and without accompanying picture,” Recommendation ITU-R BS.775-3, International Telecommunications Union, Geneva, 2012.
- [11] ATSC: “Parameterized Services Standard,” Doc. A/71:2012, Advanced Television Systems Committee, Washington, D.C., 3 December 2012.

3. COMPLIANCE NOTATION

This section defines compliance terms for use by this document:

shall – This word indicates specific provisions that are to be followed strictly (no deviation is permitted).

shall not – This phrase indicates specific provisions that are absolutely prohibited.

should – This word indicates that a certain course of action is preferred but not necessarily required.

should not – This phrase means a certain possibility or course of action is undesirable but not prohibited.

3.1 Treatment of Syntactic Elements

This document contains symbolic references to syntactic elements used in the audio, video, and transport coding subsystems. These references are typographically distinguished by the use of a different font (e.g., *restricted*), may contain the underscore character (e.g., `sequence_end_code`) and may consist of character strings that are not English words (e.g., `dynrng`).

The word “dialog” may be spelled as “dialogue” in other ATSC documents. The difference in spelling does not indicate any difference in meaning.

3.2 Symbols, Abbreviations, and Mathematical Operators

The symbols, terms, and mathematical operators used herein are as found in Section 3.4 of ATSC A/53 Part 1 [7].

4. SYSTEM OVERVIEW

As illustrated in Figure 4.1, the audio subsystem comprises the audio encoding/decoding function and resides between the audio inputs/outputs and the transport subsystem. The audio encoder(s) is (are) responsible for generating the audio elementary stream(s) which are encoded representations of the baseband audio input signals. At the receiver, the audio subsystem is responsible for decoding the audio elementary stream(s) back into baseband audio. In Figure 4.1, the arrows indicating the audio elementary stream identify the elements of the system covered by this Part.

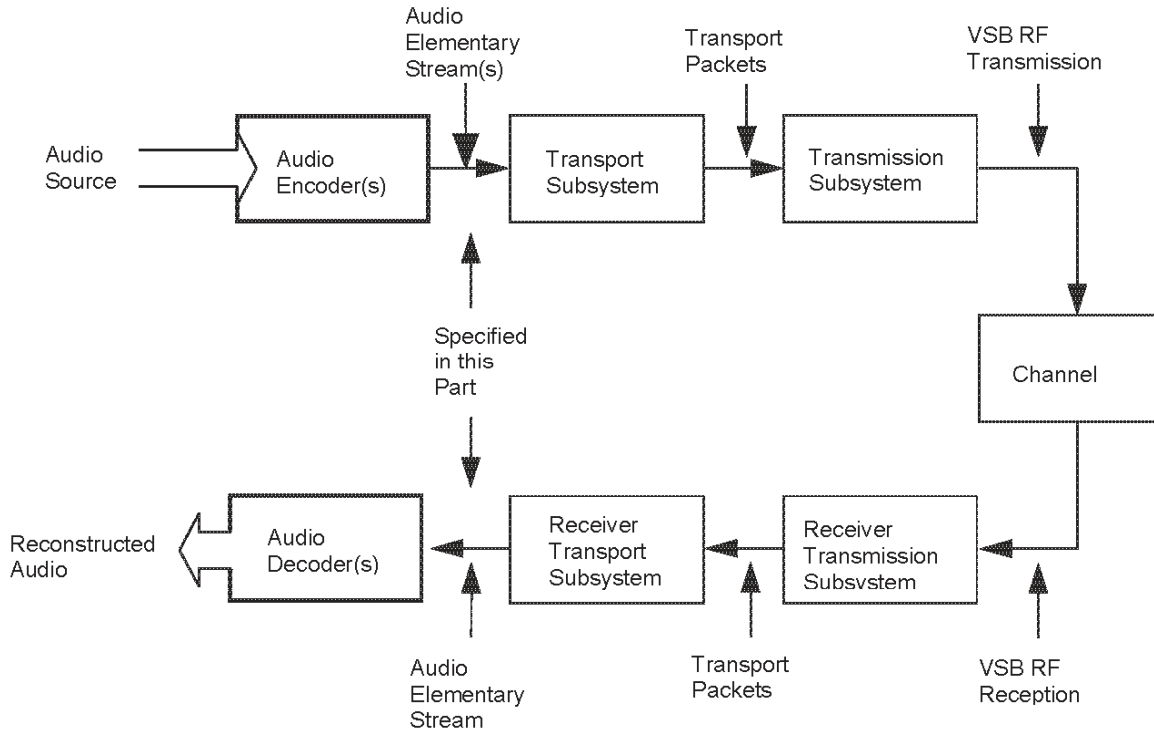


Figure 4.1 Audio subsystem in the digital television system.

5. CONSTRAINTS ON E-AC-3 CODING

This Section comprises the normative constraints for Elementary Stream (ES) components carrying audio encoded with the E-AC-3 codec. The coding of E-AC-3 audio streams shall conform to Annex E of the “Digital Audio Compression (AC-3, E-AC-3) Standard,” A/52 [1], subject to the constraints in this document.

5.1 Syntactical Element Constraints

E-AC-3 audio elementary streams shall comply with the syntax and semantics as specified in A/52 [1] including Annex E. Constraints on the coding of E-AC-3 syntactical elements shall be as constrained hereinafter.

E-AC-3 streams shall be constrained as follows:

- The value of *fscod* shall be ‘00’ (which indicates an audio sampling rate of 48 kHz).
- The value of *bsid* shall be ‘10000.’
- The value of *numblkscod* shall be ‘11’ (which indicates 6 audio blocks per frame).
- The minimum data rate of an E-AC-3 stream shall be 32 kb/s.
- The maximum data rate of an E-AC-3 stream shall be 3,024 kb/s. This maximum limit shall apply to all E-AC-3 streams regardless of the substream configurations of the streams.
- For any associated audio services carried within the same E-AC-3 stream as the main audio service, the maximum number of audio channels shall be 5.1 per associated audio service.

5.2 Sampling Rate

The system conveys digital audio sampled at a frequency of 48 kHz that shall be locked to the 27 MHz MPEG-2 system clock. The 48 kHz audio sampling clock is defined as

$$48 \text{ kHz audio sample rate} = (2 \div 1125) \times (27 \text{ MHz MPEG-2 system clock})$$

If analog signal inputs are employed, the A/D converters shall sample at 48 kHz locked to the 27 MHz clock. If digital inputs are employed, the input sampling rate shall be 48 kHz locked to the system clock, or the audio encoder shall contain sampling rate converters which convert the sampling rate to 48 kHz locked to the system clock.

5.3 Audio Coding Modes

Audio services shall be encoded using any of the audio coding modes specified in A/52 [1], with the exception of the 1+1 mode. The value of *acmod* in the E-AC-3 bit stream shall have a value in the range of 1–7, with the value 0 prohibited.

5.4 Constraints for Audio Mixing Applications

When a bitstream or independent substream is carrying an associated service that is intended to be combined with a main audio service after decoding, mixing metadata as specified in sections E2.2.2 and E2.3.1 of A/52 [1] shall be present in the bitstream or independent substream, and the following constraints shall apply to the bitstream or independent substream:

- In order to have the audio from the main audio service and the associated audio service reproduced in exact sample synchronism, it is required that all of the audio program elements be encoded frame-synchronously; i.e., if the bitstream or independent substream carrying the main audio service has sample 0 of frame *n* taken at time *t*₀, then the bitstream or independent substream carrying the associated audio service also shall have sample 0 of frame *n* taken at time *t*₀.
- The number of channels present in the associated audio service shall not exceed the number of channels present in the main audio service with which it is associated.
- An associated audio service with an audio coding mode greater than 1/0 shall not contain channel locations that are not present in the main audio service.
- The value of *mixmdate* shall be set to '1', indicating that mixing metadata is present in the stream.
- The value of *extpgmscle* shall be set to '1' when control over the global gain of the main audio service is required, indicating that *extpgmscl* data is present in the stream.
- The value of *pgmscle* shall be set to '0', indicating that *pgmscl* data is not present in the substream.
- The value of *paninfoe* may be set to '1', indicating that panning data is present in the substream, only when the audio coding mode of the substream is 1/0.
- The value of *mixdef* shall be set to '11' when control over the gain of individual channels in the main audio service is required. Otherwise the value of *mixdef* shall be set to '00'.
- The value of *mixdata2e* shall be set to '1' when additional mixing data to control the level of individual channels in the main audio service follows in the stream.
- The value of *premixcmpsel* shall be set to '0'.
- The value of *drsrc* shall be set to '0'.
- The value of *premixcmpscl* shall be set to '000'.

5.5 Dialog Level

The value of the `dialnorm` parameter in the E-AC-3 elementary bit stream shall indicate the loudness¹ of the encoded audio content (typically of the average spoken dialog) using LKFS units. LKFS and its loudness measurement algorithm are specified in ITU-R Recommendation BS.1770-3, Annex 1 [3]². (Receivers use the value of `dialnorm` to adjust the reproduced audio level to normalize the loudness.)

5.6 Dynamic Range Compression

Dynamic Range Compression (DRC) is an important feature of the E-AC-3 codec and is documented at length in ATSC A/85 [8] Section 9 and Annex F.

5.6.1 Dynamic Range Compression – “Line Mode”

Each encoded audio block contains a dynamic range control word (`dynrng`) that is used by decoders (by default) to alter the level of the reproduced audio. The control words allow the decoded signal level to be increased or decreased by up to 24 dB. In general, elementary streams may have dynamic range control words inserted or modified without affecting the encoded audio. When it is necessary to alter the dynamic range of audio programs that are broadcast, the dynamic range control word should be used.

5.6.2 Dynamic Range Compression – “RF Mode”

Each encoded audio frame contains a dynamic range control word (`compr`) that may be optionally used by decoders to render the audio with a very narrow dynamic range. The control words allow the decoded signal level to be increased or decreased by up to 48 dB.

6. INTRODUCTION TO E-AC-3 AUDIO SERVICES

As defined in A/52 [1], an E-AC-3 elementary stream contains the encoded representation of one or more audio services. Multiple E-AC-3 elementary streams may be provided. There are a number of audio service types that may be coded into each elementary stream. There is a *complete main service* and there are various types of *associated services*.

Associated services can contain complete program mixes containing all audio program elements (dialog, music, effects, etc.) that are intended to be presented to a listener. This is indicated by the `full_service_flag` bit in the `E-AC-3_audio_stream_descriptor()` being set to a value of ‘1’ (see A/53-3 [5] and A/52, Annex G [1]). Alternatively, associated services can be combined with other audio services to provide enhanced functionality, such as VI, HI, and ME.

6.1 Options for Providing Video Description

The term “Video Description” is used for to improve the experience of the visually impaired, such as the insertion of audio narrated descriptions of a television program’s key visual elements into natural pauses between the program’s dialog in statute and regulation. The abbreviation “VI” has been retained as that is the label for this value of the `bsmod` field in A/52 [1].

¹ Methods to measure loudness and different modes for `dialnorm` management are explained in the ATSC Recommended Practice A/85 “Techniques for Establishing and Maintaining Audio Loudness for Digital Television” [8]; see particularly Section 5.2.

² A `dialnorm` value based on an “A” weighted integrated measurement (LAeq) (ANSI S1.4) [2] prior to the publication of this document need not be re-measured with BS.1770 [3].

7. CONSTRAINTS ASSOCIATED WITH SPECIFIC AUDIO SERVICE TYPES

This section provides constraints on specific audio service types when used in the ATSC system. Additional audio service types are defined in A/52 [1].

7.1 For service_type Values of 2 and 3

The following subsections document the constraints for adding an E-AC-3 audio component to services having service_type values 2 or 3. The signaling of audio components is found in Part 3 [5]. For a main audio service, the maximum number of audio channels shall be 5.1.

Note: Services having service_type values 2 or 3 are required by Part 3 [5] to contain an AC-3 CM audio service. Multiple CM services may be signaled.

7.1.1 Complete Main (CM) Audio Service

The CM type of audio service is signaled by using bsmo_d value ‘000’ as described in A/52 [1]. The CM type of audio service shall contain a complete audio program (including dialog, music, and effects). The CM service may contain from 1 to 5.1 audio channels.

7.1.2 Main Audio Service, Music and Effects (ME)

The ME type of audio service is signaled by using bsmo_d value ‘001’ as described in A/52 [1]. The ME type of audio service is prohibited for these service types.

7.1.3 Visually Impaired (VI)

The VI type of audio service is signaled by using bsmo_d value ‘010’ as described in A/52 [1]. The VI type of audio service shall be constrained to be a complete program mix containing aspects designed to improve the experience of the visually impaired, such as the insertion of audio narrated descriptions of a television program’s key visual elements into natural pauses between the program’s dialog³. The VI service may be coded using any number of channels (up to 5.1).

7.1.4 Hearing Impaired (HI)

The HI type of audio service is signaled by using bsmo_d value ‘011’ as described in A/52 [1] and shall be a program mix with enhanced intelligibility. The HI service may be coded using any number of channels (up to 5.1).

7.1.5 Additional bitstream Constraints

The E-AC-3 stream shall consist of a single independent substream.

An independent substream consists of a sequence of E-AC-3 synchronization frames. These synchronization frames shall comply with the following constraints:

strmtyp: This field shall be set to ‘00’ to indicate an independent substream.

substreamid: This field shall be set to ‘000’.

7.2 For service_type Value 7

The following subsections document the constraints for adding an E-AC-3 main audio component for services having service_type value 7. For a main audio service, the maximum number of audio channels shall be 7.1. Zero or more audio services are permissible.

The Stream info details shall be as defined in Section 7.2.9.

³ The term “Video Description” is sometimes used to describe such services.

7.2.1 Complete Main (CM) Audio Service

The CM type of audio service is signaled by using bsmod value ‘000’ as described in A/52 [1]. The CM type of main audio service shall contain a complete audio program (including dialog, music, and effects). This is the type of audio service normally provided. A/52 [1] establishes the capability to send from 1 to 7.1 audio channels within the CM service. A CM service can be further enhanced by means of the associated services described in Section 7.2.

7.2.2 Main Audio Service, Music and Effects (ME)

The ME type of audio service is signaled by using bsmod value ‘001’ as described in A/52 [1]. The ME type of main audio service shall contain the music and effects portion of an audio program. The ME service may contain from 1 to 7.1 audio channels. The primary program dialog shall not be included in a ME service. See the D type of associated service for more information.

7.2.3 Visually Impaired (VI)

The VI type of audio service is signaled by using bsmod value ‘010’ and shall be a program mix designed to improve the experience of the visually impaired, such as the insertion of audio narrated descriptions of a television program’s key visual elements into natural pauses between the program’s dialog. Per A/52 [1], the VI associated service is either a complete program mix containing music, effects, dialog, and additionally a narration that describes the picture content (full_service_flag = ‘1’) or an additional associated narration that describes the picture content (full_service_flag = ‘0’). A VI service that is a complete program mix may be coded using any number of channels (up to 7.1). A VI service that is an additional associated narration may be coded using any number of channels up to 5.1.

Note: See Section 8 for further details of signaling this service type.

7.2.4 Hearing Impaired (HI)

The HI type of audio service is signaled by using bsmod value ‘011’ as described in A/52 [1] and typically is a complete program mix with enhanced intelligibility. The HI service may be coded using any number of channels (up to 7.1).

7.2.5 Dialog (D)

The D type of audio service is signaled by using bsmod value ‘100’ as described in A/52 [1] and contains program dialog intended for use with an ME main audio service. The language of the D service is indicated in the E-AC-3_audio_descriptor(). A complete audio program is formed by simultaneously decoding a D and an ME service and mixing the audio from the D service into the audio from the ME service. Mixing metadata is found in the D service bitstream or substream. Multiple D services (e.g., in different languages) may be associated with a single ME service. A D service may be coded using any number of channels up to 5.1.

7.2.6 Commentary (C)

The C type of audio service is signaled by using bsmod value ‘101’ as described in A/52 [1]. The C associated service is a complete program mix containing music, effects, dialog, and some special commentary. This service may be provided using any number of channels (up to 7.1).

7.2.7 Voice-Over (VO)

The VO type of audio service is signaled by using bsmod value ‘111’ as described in A/52 [1]. The VO associated service is a monophonic service. The intended use is to add alternative dialog to the program.

7.2.8 Additional Bitstream Constraints

The E-AC-3 stream shall contain no more than four independent substreams. An E-AC-3 stream shall contain no more than one dependent substream, and this dependent substream shall be associated with independent substream 0 to support audio services of up to 7.1 channels. For associated audio services carried in independent substreams within the same E-AC-3 stream as the main audio service, the maximum number of audio channels shall be 5.1 per associated audio service.

7.2.8.1 Independent Substream Constraints

An independent substream consists of a sequence of E-AC-3 synchronization frames. These synchronization frames shall comply with the following constraints:

strmtyp: This field shall be set to '00' to indicate an independent substream.

substreamid: This field shall be set to '000', '001', '010', or '011'.

7.2.8.2 Dependent Substream Constraints

The dependent substream consists of a sequence of E-AC-3 synchronization frames. These synchronization frames shall comply with the following constraints:

strmtyp: This field shall be set to '01' to indicate a dependent substream.

substreamid: This field shall be set to '000'.

7.2.8.3 Substream Configuration for Delivery of an Audio Service of up to 5.1 Channels

To deliver a main audio service containing up to 5.1 channels, there shall be no dependent (strmtyp = '01') substream 0 present in the E-AC-3 audio bitstream. Furthermore:

- Per Section 5.3, the value of acmod shall be between 1 and 7.
- The value of ifeon may be '1'.
- There may be present independent (strmtyp = '00') substreams in addition to independent substream 0.

7.2.8.4 Substream Configuration for Delivery of an Audio Service of More than 5.1 Channels

To deliver a main audio service containing more than 5.1 channels, an independent (strmtyp = '00') substream with an acmod value of '111' and a single dependent (strmtyp = '01') substream with a substreamid value of 0 shall be included in the E-AC-3 stream as specified by A/52 [1] Section E2.3.1.2.

The channel configuration of the complete elementary stream shall be constrained to be no more than 7.1 channels defined by the combination of the acmod field (value of '111') carried in the independent substream along with the acmod (value of '010') and chanmap fields carried in the dependent substream. Further, the value of ifeon shall be '1'. The loudspeaker locations supported by E-AC-3 are defined in [8].

The following rules apply to channel numbers and substream use:

- When more than 5.1 channels of audio are to be delivered, independent substream 0 of an E-AC-3 stream shall be configured as a downmix of the complete program.
- Additional channels necessary to deliver up to 7.1 channels of audio shall be carried in dependent substream 0.
- When up to 5.1 channels of decoded output are required, a receiver shall decode independent substream 0 and shall ignore dependent substream 0.
- When more than 5.1 channels of decoded output are required (up to 7.1 channels), a receiver shall decode both independent substream 0 and dependent substream 0.

- Support for decoding of dependent substream 0 is optional.

7.2.9 Stream Info Details for stream_type 0x87, service_type value 7

When a stream of stream type 0x87 is present in a virtual channel of service_type value 7 the component_list_descriptor() per ATSC A/71 [11] shall be present. The field length_of_details in the component_list_descriptor() is required by A/71 [11] to contain the count of bytes that are in the stream_info_details() portion of this component_list_descriptor(). The stream_info_details() portion shall be constructed as shown in Table 7.1, with the contents of the fields as defined below:

Table 7.1 stream_info_details() Syntax for Stream Type 0x87 (E AC-3)

Syntax	No. of Bits	Format
stream_info_details() {		
fscod	2	uimsbf
sourcefscod	1	uimsbf
reserved	1	'1'
max_channels	4	uimsbf
future_fields()	var	
}		

fscod – This field shall be set as required by A/52 [1] Section E2.3.1.4.

sourcefscod – This field shall be set as required by A/52 [1] Section E2.3.1.63.

max_channels – This field shall indicate the maximum number of audio channels which might be carried by this service.

future_fields() – This variable-length data structure is established to indicate the explicit ability to convey additional information in future versions of this standard. The value in the length_of_details which immediately precedes this instance of stream_info_details() in the component_list_descriptor() (see A/71 [11]) includes the count of the bytes in this field. In the present standard, the value of length_of_details is 1.

Note: Per A/71 [11], if an unsupported value of length_of_details is encountered, receivers are expected to conclude that the stream is not supported for decoding.

Note: Per A/71 [11], this structure will be placed in the Component List Descriptor within either a Terrestrial Virtual Channel Table or a Cable Virtual Channel Table.

8. MIXING OF MAIN AND ASSOCIATED AUDIO SERVICES

8.1 Introduction

This section describes the decoding and mixing of Main and Associated Audio services, including the usage of mixing metadata parameters from the E-AC-3 stream, and considerations for system designers. Support for dual stream decoding is optional and is only required for applications that require mixing of Main and Associated audio services.

8.2 Mixing Metadata

A/52 Section E3.10 [1] describes the use of metadata to control program mixing. The following subsections provide additional guidance.

8.2.1 Control Smoothing

Mixing metadata is delivered once per E-AC-3 frame. To prevent audible glitches during fast gain changes, interpolation between successive gain values should be performed in the receiver.

8.2.2 Output Limiting

When audio services are decoded and combined (“mixed”) to produce a single signal, the mixing process can result in signal overload conditions that significantly reduce audio quality. Two common approaches to avoiding overload conditions are: 1) to attenuate the inputs before mixing such that the mixing does not result in overload, or 2) to use a compressor/limiter. The use of a compressor/limiter is recommended, as it results in consistent output level.

8.3 Examples of Decoding and Mixing Configurations

This section describes three example configurations of the audio decoding and mixing chain in a receiver, together with examples of the E-AC-3 descriptor contents that would be used to signal the combination of audio services in each example. Block diagrams of each receiver configuration are shown. For the sake of simplicity, the following control paths and other components are not shown in these figures:

- The routing of the AC-3 bitstream directly to a suitable digital output interface (e.g. S/PDIF and/or HDMI), bypassing the audio decoder and mixing chain, in situations where two audio PIDs are transmitted.
- Conversion of the first independent substream (I0) of the bitstream carrying the main audio service to AC-3 (as described in section 8), and routing of this AC-3 output stream directly to a suitable digital audio output interface (S/PDIF and/or HDMI), bypassing the decoding and mixing chain.
- Passing of the downmix scaling parameters from the Main audio bitstream (e.g. `ltrcmixlev`, `ltrtsurmixlev`) to the post-mixer downmix block to ensure that downmix balance is maintained at the output of the mixer.
- User controls to adjust the balance between main audio and associated audio.

8.3.1 Mixing of AC-3 Main with E-AC-3 Associated Audio

In this example, a 5.1-channel Complete Main (CM) audio service carried in an AC-3 bitstream is mixed with a mono associated Video Description (VI) audio service carried in an E-AC-3 bitstream. The configuration of the receiver's audio decoding and mixing components is shown in Figure 8.1.

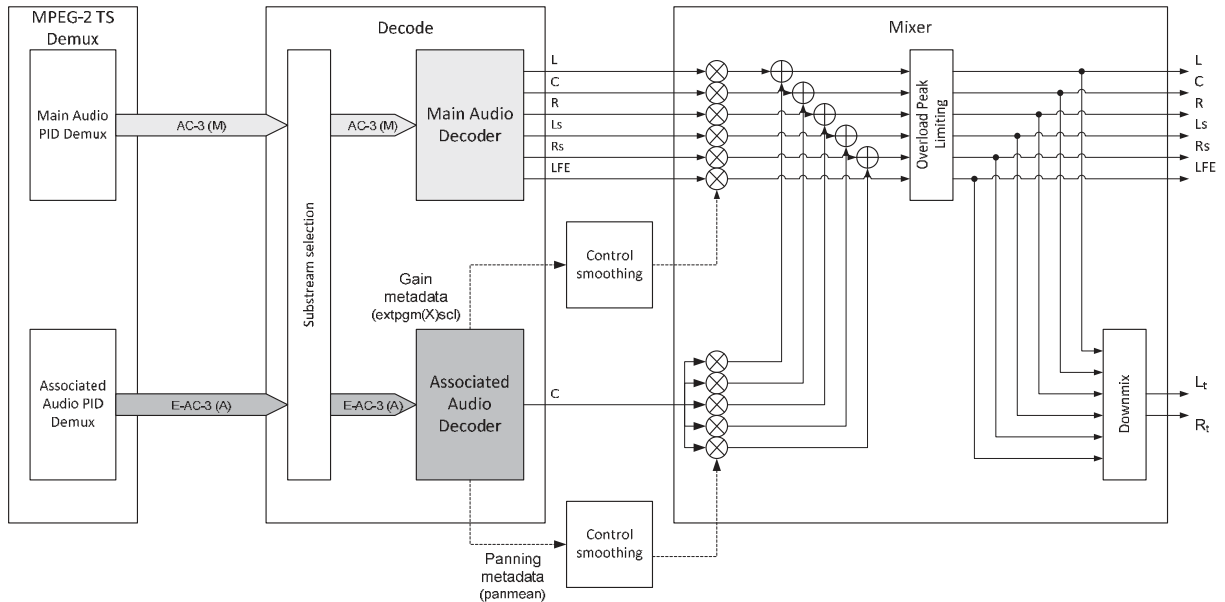


Figure 8.1 Mixing an AC-3 main with an E-AC-3 associated service.

Control of the balance between the main and associated audio services is provided by gain metadata parameters delivered in the E-AC-3 associated audio bitstream. Additionally, the mono associated audio service can be panned (shown above by the Associated Audio Decoder 5-way output split) to any of the full range channel locations, using panning metadata parameters delivered in the E-AC-3 associated audio bitstream.

The diagram also shows dual outputs from the mixer: One is a 5.1 channel output suitable for delivery to a re-encoding process for output over S/PDIF or HDMI, and the second is a 2-channel decoded output that is generated by downmixing the output of the audio mixer.

The configuration of each audio service is described in its respective audio descriptor present in the PMT: an AC-3 audio descriptor to describe the main audio service and an E-AC-3 audio descriptor to describe the associated audio service. The *mainid* and *asvc* fields in those respective descriptors are set to values that establish the linkage for the services. In this example, a *mainid* (AC-3) value of '000' and an *asvc* (E-AC-3) value of '00000001' are used, indicating that the services should be demultiplexed and decoded together. The configuration of the AC-3 audio descriptor for the main audio service is shown in Table 8.1. The configuration of the E-AC-3 audio descriptor for the associated audio service is shown in Table 8.2.

Table 8.1 AC-3_audio_stream_descriptor Configuration – Main Audio Service

Parameter	No. of Bits	Value	Meaning
descriptor_tag	8	0x81	Assigned value for AC-3 descriptor
descriptor_length	8	0x09	9 bytes
sample_rate_code	3	'000'	48 kHz
bsid	5	'00110'	bsid 6
bit_rate_code	6	'001110'	384 kbit/s
surround_mode	2	'00'	Not indicated
bsmod	3	'000'	Complete Main
num_channels	4	'0111'	acmod 3/2
full_svc	1	'1'	Full Service
langcod	8	0x00	No longer used
mainid	3	'000'	Main audio service ID 0
priority	2	'11'	Priority not specified
reserved	3	'000'	
language_flag	1	'1'	Language field present
language_flag_2	1	'0'	Language2 field not present
reserved	6	'111111'	
language	3*8	0x656E67	eng (English)

Table 8.2 E-AC-3 audio stream descriptor Configuration – Associated VI Audio Service

Parameter	No. of Bits	Value	Meaning
descriptor_tag	8	0xCC	Assigned value for E-AC-3 descriptor
descriptor_length	8	0x07	7 bytes
reserved	1	'1'	
bsid_flag	1	'1'	bsid field present
mainid_flag	1	'0'	mainid field not present
asvc_flag	1	'1'	asvc field present
mixinfoexists	1	'1'	mixing metadata exists in independent substream 0
substream1_flag	1	'0'	substream1 field not present
substream2_flag	1	'0'	substream2 field not present
substream3_flag	1	'0'	substream3 field not present
reserved	1	'1'	
full_service_flag	1	'0'	Substream I0 does not contain a full audio service
audio_service_type	3	'010'	Video Description (VI) service
number_of_channels	3	'000'	Mono audio
language_flag	1	'1'	Language field present
language_flag_2	1	'0'	Language2 field not present
reserved	1	'1'	
bsid	5	'10000'	bsid 16
asvc	8	'00000001'	Linked to Main Service with mainid value of 0
language	3*8	0x656E67	eng (English)

8.3.2 Mixing of main and associated services in separate bitstreams

In this example, the 5.1-channel Complete Main (CM) audio service carried in the E-AC-3 primary bitstream is decoded as 2-channels (L_M, R_M), and then mixed with a mono associated (M_A)

Voiceover (VO) audio service carried in a separate E-AC-3 bitstream. The configuration of the receiver's audio decoding and mixing components is shown in Figure 8.2.

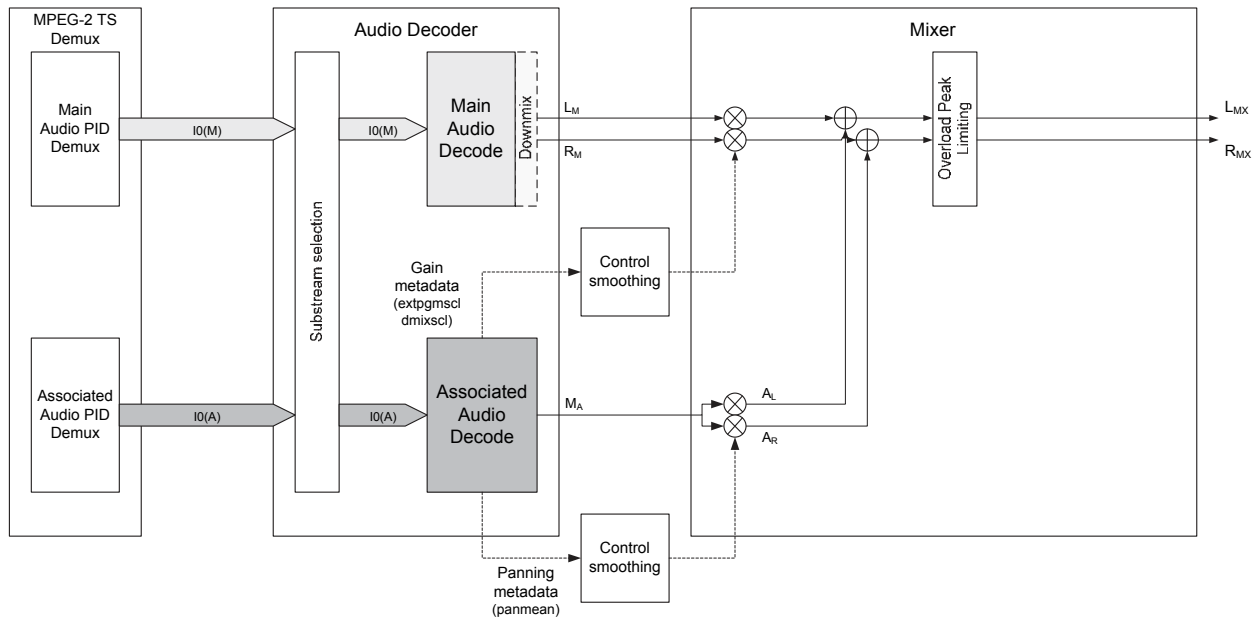


Figure 8.2 Mixing of a downmixed multichannel main with mono associated audio

The balance between the main and associated audio service is primarily controlled using the `extpgmscl` gain parameter. The `dmixscl` parameter is used to apply additional scaling to the downmixed audio output from the main audio decoder.

Additionally, the mono associated audio service can be panned between the left and right output channels using the `panmean` parameter.

The configuration of each audio service is described in its respective E-AC-3 audio descriptor present in the PMT: one descriptor for the main audio service and another for the associated audio service. The `mainid` and `asvc` fields in those respective descriptors are set to values that establish the linkage for the services. In this example, a `mainid` value of '001' and an `asvc` value of '00000010' are used, indicating that the services should be demultiplexed and decoded together. The configuration of the E-AC-3 audio descriptor for the main audio service is shown in Table 8.3. The configuration of the E-AC-3 audio descriptor for the associated VO audio service is shown in Table 8.4.

Table 8.3 E-AC-3 audio stream descriptor Configuration – Main Audio Service

Syntax	No. of Bits	Value	Meaning
descriptor_tag	8	0xCC	Assigned value for E-AC-3 descriptor
descriptor_length	8	0x07	7 bytes
reserved	1	'1'	
bsid_flag	1	'1'	bsid field present
mainid_flag	1	'1'	mainid field present
asvc_flag	1	'0'	asvc field not present
mixinfoexists	1	'0'	mixing metadata not present in independent substream 0
substream1_flag	1	'0'	substream1 field not present
substream2_flag	1	'0'	substream2 field not present
substream3_flag	1	'0'	substream3 field not present
reserved	1	'1'	
full_service_flag	1	'1'	Substream I0 contains a full service
audio_service_type	3	'000'	Complete Main
number_of_channels	3	'100'	Multichannel Audio (> 2-channel <= 3/2 + LFE)
language_flag	1	'1'	language field present
language_flag_2	1	'0'	language2 field not present
reserved	1	'1'	
bsid	5	'10000'	bsid 16
reserved	3	'111'	
priority	2	'11'	Priority not specified
mainid	3	'001'	Main audio service ID 1
language	3*8	0x656E67	eng (English)

Table 8.4 E-AC-3 audio stream descriptor Configuration – Associated VO audio service

Parameter	No. of Bits	Value	Meaning
descriptor_tag	8	0xCC	Assigned value for E-AC-3 descriptor
descriptor_length	8	0x07	7 bytes
reserved	1	'1'	
bsid_flag	1	'1'	bsid field present
mainid_flag	1	'0'	mainid field not present
asvc_flag	1	'1'	asvc field present
mixinfoexists	1	'1'	mixing metadata exists in substream I0
substream1_flag	1	'0'	substream1 field not present
substream2_flag	1	'0'	substream2 field not present
substream3_flag	1	'0'	substream3 field not present
reserved	1	'1'	
full_service_flag	1	'0'	Substream I0 does not contain a full audio service
audio_service_type	3	'111'	Voiceover (VO) service
number_of_channels	3	'000'	Mono (acmod 1/0)
language_flag	1	'1'	Language field present
language_flag_2	1	'0'	Language2 field not present
reserved	1	'1'	
bsid	5	'10000'	bsid 16
asvc	8	'00000010'	Linked to Main Service with mainid value of 1
language	3*8	0x656E67	eng (English)

8.3.3 Mixing of 5.1-channel main with 3-channel associated audio

In this example, the 5.1-channel downmix of a 7.1-channel Music and Effects (ME) audio service is mixed with an associated French language Dialog (D) audio service. The Music and Effects audio service is carried in independent substream 0 and dependent substream 0 of the E-AC-3 bitstream. Independent substream 0 is decoded to produce the 5.1-channel downmix of the overall Music and Effects service. Three Dialog audio services (English, French and Spanish) also are carried in independent substreams 1, 2 and 3 of the same E-AC-3 bitstream, respectively. In this example, the French Dialog service, which consists of 3 channels (LCR) in independent substream 2, is selected for decoding and mixing with the Music and Effects service from independent substream 0. The configuration of the receiver's audio decoding and mixing components is shown in Figure 8.3.

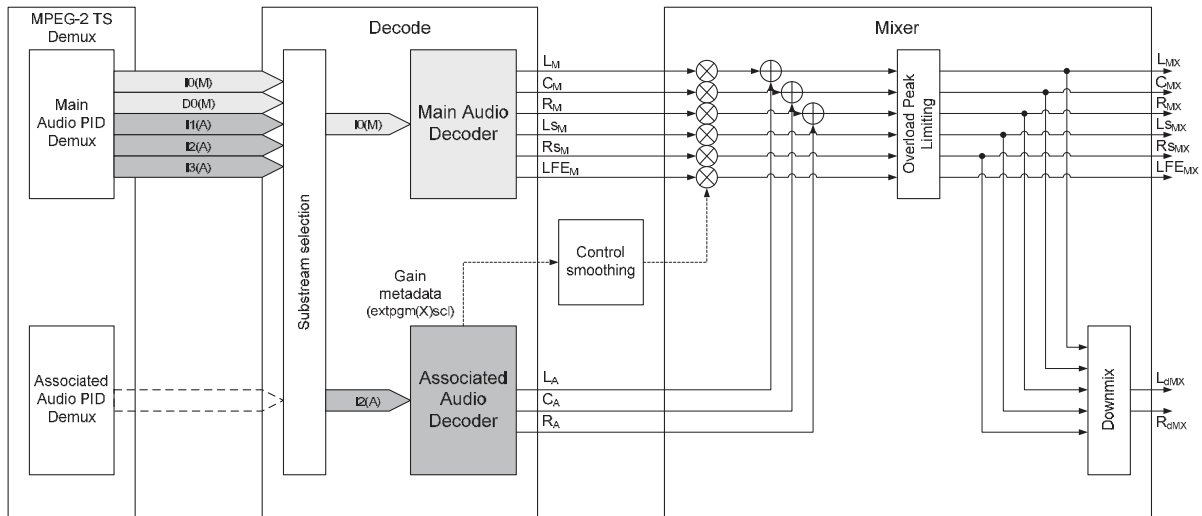


Figure 8.3 Mixing a 5.1-channel Main with a 3-channel Associated.

While the balance between the Dialog and the Music and Effects audio services is primarily defined before encoding, adjustments to this balance can be achieved by using the `extpgmscl` and `extpgm(X)sc1`⁴ parameters carried in independent substream 2. If the `extpgmaux1sc1` and `extpgmaux2sc1` parameters are present in independent substream 2, they are not used, as the additional two channels of the original 7.1-channel Music and Effects audio service are not being decoded.

The configuration of all audio services is described in a single E-AC-3 audio descriptor present in the PMT. The main audio service is described in the main body of the descriptor, and the associated audio services carried in independent substreams 1, 2, and 3 are described in the `substream1`, `substream2`, and `substream3` fields, respectively. Additionally, as each Dialog service carries a different language, the language of each associated service is described using the `substreamX_lang` field for each independent substream. The configuration of the E-AC-3 audio descriptor is shown in Table 8.5.

⁴ Where “X” can be L, C, R, Ls, Rs, LFE, aux1, aux2. See A/52 [1] Section E3.10.6.

Table 8.5 E-AC-3_audio_stream_descriptor Configuration – ME and associated D audio services

Syntax	No. of Bits	Value	Meaning
descriptor_tag	8	0xCC	Assigned value for E-AC-3 descriptor
descriptor_length	8	0x13	19 bytes
reserved	1	'1'	
bsid_flag	1	'1'	bsid field is present
mainid_flag	1	'1'	mainid field is present
asvc_flag	1	'0'	asvc field is not present
mixinfoexists	1	'0'	Mixing metadata is not present in independent substream 0
substream1_flag	1	'1'	substream1 field is present
substream2_flag	1	'1'	substream2 field is present
substream3_flag	1	'1'	substream3 field is present
reserved	1	'1'	
full_service_flag	1	'1'	independent substream 0 carries a full service
audio_service_type	3	'001'	Music and Effects (ME)
number_of_channels	3	'101'	Multichannel Audio > 3/2+LFE channels
language_flag	1	'0'	language field is not present
language_flag_2	1	'0'	language2 field is not present
reserved	1	'1'	
bsid	5	'10000	bsid 16
reserved	3	'111'	
priority	2	'11'	Priority not indicated
mainid	3	'000'	Main audio service ID 0
/*substream1*/			
reserved	1	'1'	
substream_priority	1	'0'	Not specified
audio_service_type	3	'100'	Dialog
number_of_channels	3	'100'	Multichannel Audio (> 2-channel <= 3/2 + LFE)
/*substream2*/			
reserved	1	'1'	
substream_priority	1	'0'	Not specified
audio_service_type	3	'100'	Dialog
number_of_channels	3	'100'	Multichannel Audio (> 2-channel <= 3/2 + LFE)
/*substream3*/			
reserved	1	'1'	
substream_priority	1	'0'	Not specified
audio_service_type	3	'100'	Dialog
number_of_channels	3	'100'	Multichannel Audio (> 2-channel <= 3/2 + LFE)
substream1_lang	3*8	0x656E67	eng (English)
substream2_lang	3*8	0x667261	fra (French)
substream3_lang	3*8	0x737061	spa (Spanish)

9. CONVERSION OF E-AC-3 INDEPENDENT SUBSTREAM 0 TO AC-3

When a device contains an E-AC-3 decoder and has a digital output intended to feed the input of a downstream decoder that is only AC-3 capable, then conversion of the E-AC-3 stream to an AC-3 stream is necessary.

This section describes, in general terms, the recommended process for converting an E-AC-3 stream to an AC-3 stream. This conversion process can be performed using only the audio present in independent substream 0 in the E-AC-3 stream. Any additional dependent and independent substreams in the E-AC-3 stream thus can be ignored by the conversion process. Figure 9.1 shows a diagram of the recommended conversion process.

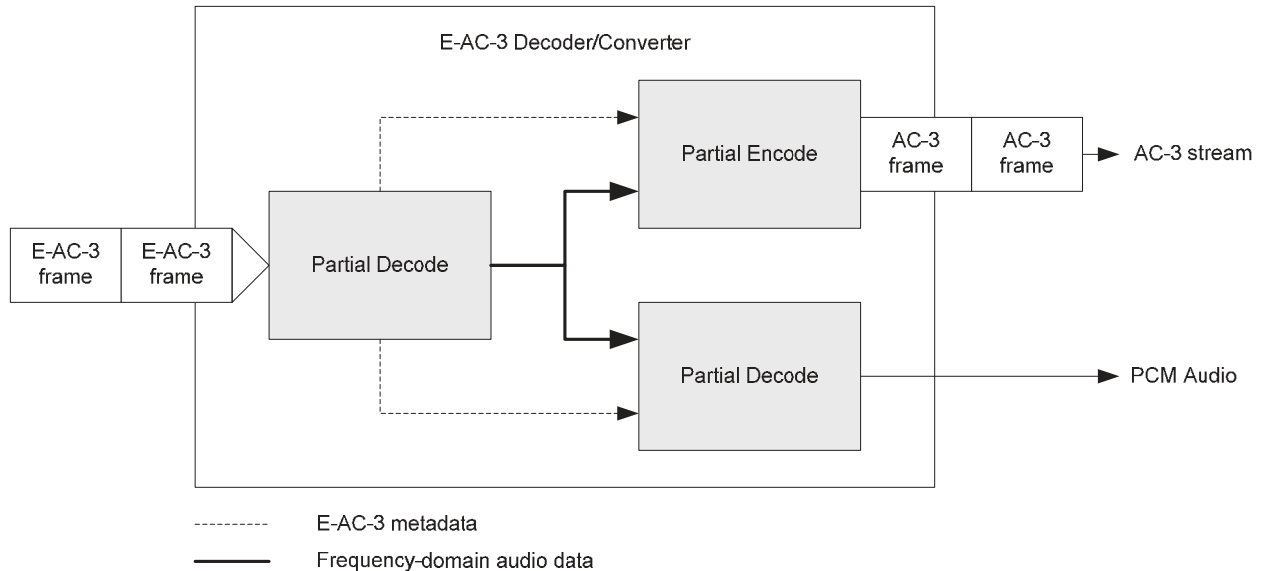


Figure 9.1 E-AC-3 to AC-3 conversion process.

As E-AC-3 is a superset of AC-3, sharing much of the bitstream syntax and audio coding tools of AC-3, it is possible to perform the conversion process without completely decoding the E-AC-3 stream to PCM, with subsequent re-encoding of the PCM to AC-3. Rather, a partial decode process is performed on each E-AC-3 frame. This process uses both frequency-domain audio data and bitstream metadata from the E-AC-3 frame. The frequency domain audio data is passed to a partial re-encode process, which is used to create an AC-3 frame. The E-AC-3 metadata is used to create equivalent AC-3 metadata (including reformatting to match the syntax and frame alignment of AC-3), which is passed to the partial re-encode process. The partial re-encode process adds the metadata to the AC-3 audio blocks and outputs a complete AC-3 frame with the required six AC-3 audio blocks.

To facilitate a minimum-complexity conversion process, the E-AC-3 bitstream includes pre-calculated AC-3 bit allocation data for converted AC-3 frames. Consequently, a bit allocation calculation is not needed in the conversion process.

As illustrated in figure 9.1, the frequency domain audio data produced by the first E-AC-3 partial decode process is then passed to a subsequent E-AC-3 final decode process operating in parallel with the AC-3 partial encode process. This final E-AC-3 decode process converts the frequency domain audio data to PCM. The dual path configuration enables both PCM and AC-3 outputs to be active when processing an E-AC-3 stream.

10. ENCODING AND DECODING OF 7.1-CHANNEL PROGRAMS

10.1 Introduction

This section describes the delivery of 7.1-channel programs using E-AC-3, including the encoding process, the structure of the encoded 7.1-channel bitstream, and the decoding process.

Figure 10.1 shows an example of the channel configuration of a typical 7.1-channel program, with left rear and right rear surround channels supplementing the conventional 5.1-channel layout. It should be noted that multiple 7.1-channel program configurations can be derived from channels labeled per SMPTE ST 428-3 [9]. E-AC-3 is capable of supporting such configurations in both encoders and decoders. Note: ITU-R BS.775-3 [10] also has helpful information regarding speaker positioning and labeling.

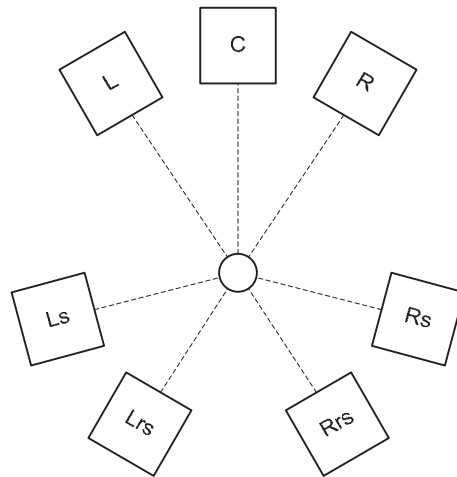


Figure 10.1 Example 7.1-channel program configuration.

10.2 Encoding Process

Figure 10.2 shows an example of the structure of a 7.1-channel E-AC-3 encoder. In the figures following, frames of independent substream 0 are denoted as “I0,” and frames of dependent substream 0 are denoted as “D0.”

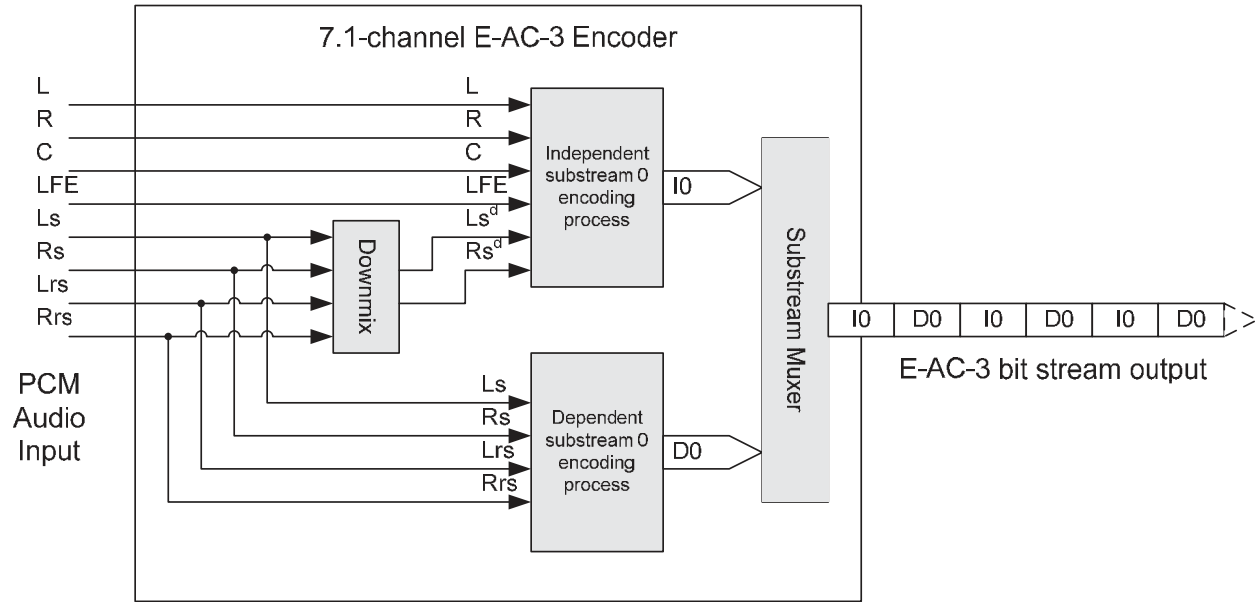


Figure 10.2 7.1 channel E-AC-3 encoder.

Eight channels of PCM audio are input to the E-AC-3 encoder. The encoder creates a 5.1-channel downmix⁵ from the 7.1-channel input. For the channel configuration shown in Figure 10.1, the four surround channels (Ls, Rs, Lrs, Rrs) of the 7.1-channel program are downmixed into two surround channels (Ls^d and Rs^d), and are combined with the original L, R, C and LFE channels to form the 5.1-channel downmix. The 5.1-channel downmix is input to a substream encoding process, which outputs a sequence of E-AC-3 frames with a *strmtyp* value of 0 (independent), and a *substreamid* value of 0 (I0). The four original surround channels are fed to a second substream encoding process, which outputs a sequence of E-AC-3 frames with a *strmtyp* value of 1 (dependent), and a *substreamid* value of 0 (D0).

The substream multiplexer time-multiplexes frames output by the substream encoding processes to form the complete E-AC-3 bit stream.

10.3 Bitstream Structure

E-AC-3 uses a scalable substream structure to deliver programs containing more than 5.1 channels of audio. The use of substreams, and the ordering of the frames of each substream within the E-AC-3 bit stream, is specified in ATSC A/52 [1] Section E2.3.1.2. For a 7.1-channel program, two substreams are used.

Figure 10.3 shows the structure of the E-AC-3 bit stream that required by Section 7.2.8.3 and illustrated by the example in Section 10.2. In the following figures, “(M)” denotes main audio service“(A)” denotes associated services. The stream begins with one frame from independent substream 0, followed by one frame from dependent substream 0. The frame from dependent substream 0 is dependent on the frame from independent substream 0, such that the frame from dependent substream 0 cannot be decoded without also decoding the preceding frame in the bit stream from independent substream 0.

⁵ The algorithm for the downmix is an implementation choice.

Figure 10.3, then, illustrates the frame structure of an E-AC-3 stream delivering a 7.1-channel main audio program carried in independent substream 0 and dependent substream 0.

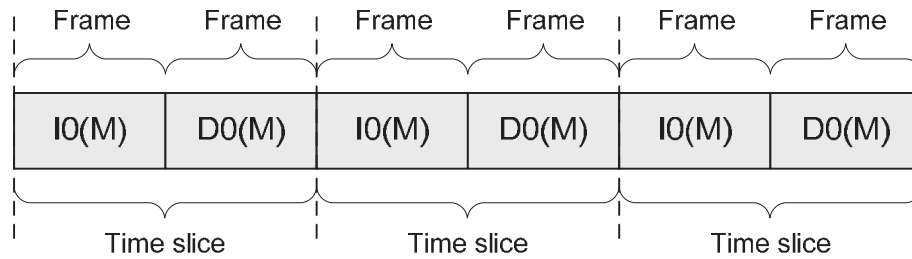


Figure 10.3 Fragment of an E-AC-3 stream structure comprising a 7.1-channel Main Audio Service.

If the bitstream also contains associated audio programs they are encoded in additional independent substreams. Figure 10.4 illustrates that one frame from each additional independent substream follows each dependent substream frame. The figure shows the independent substream 1, “I1”, following the dependent substream from the main program “D0(M)”. Figure 10.4, then, illustrates the structure of an E-AC-3 stream that is delivering a 7.1-channel main audio program carried in independent substream 0 and carried in dependent substream 0 and an associated audio program carried in independent substream 1.

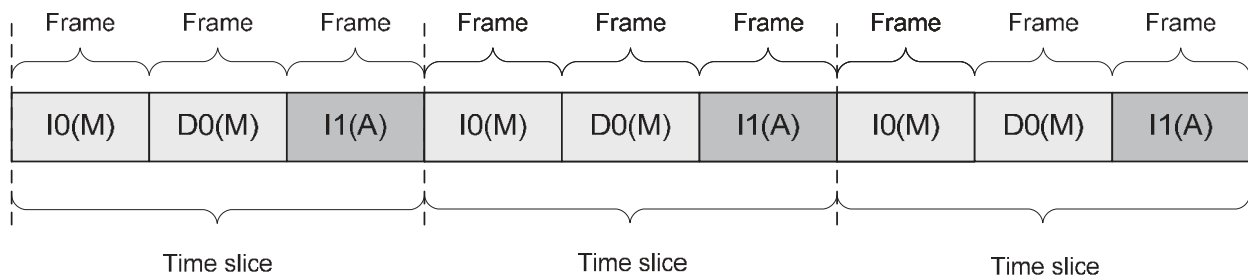


Figure 10.4 Fragment of an E-AC-3 stream structure comprising a 7.1-channel Main Audio Service and a 2-channel Associated Audio Service.

If the bitstream also contains associated audio services, they are encoded in additional independent substreams. Figure 10.4 illustrates that one frame from each additional independent substream follows each dependent substream frame. The figure shows independent substream 1, “I1”, following the dependent substream from the main service “D0(M)”. Since a 7.1-channel main audio service (as described in Section 10.2) consists of a 5.1 channel downmix (carried in independent substream 0) and the four surround channels (carried in dependent substream 0), Figure 10.4 illustrates the structure of such an E-AC-3 stream delivering a 7.1-channel main audio service and an associated audio service carried in independent substream 1.

10.4 Decoding Processes

10.4.1 7.1 Decode Process

When decoding streams encoded with 7.1 audio channels, the decoder will typically output 7.1 audio channels or fewer. To reproduce 7.1 decoded audio channels, both independent substream 0 and dependent substream 0 are decoded simultaneously. As described in Section 10.2, the

independent substream contains a 5.1-channel downmix of the 7.1-channel audio content. It therefore is necessary to replace the two surround channels from the downmixed 5.1 channel audio with the four discrete surround channels from the original 7.1 channel mix. The decoder compares the channel locations present in independent substream 0 with those in dependent substream 0. When Ls and Rs channels are found to be present in both substreams, the decoder selects for output the Ls and Rs channels from the dependent substream, as these are the original Ls and Rs channels from the 7.1-channel program mix. The decoding method is illustrated in Figure 10.5, the arrangement of which effectively reverses the encoding process outlined in Section 10.2.

While it might be possible to use a method that “undoes” the downmix in the decoder, the E-AC-3 stream instead carries replacement channels in the dependent substream to overwrite the channels in the independent substream that contain downmixed audio. This method avoids audible coding artifacts that can be caused by techniques to “reverse” the downmix.

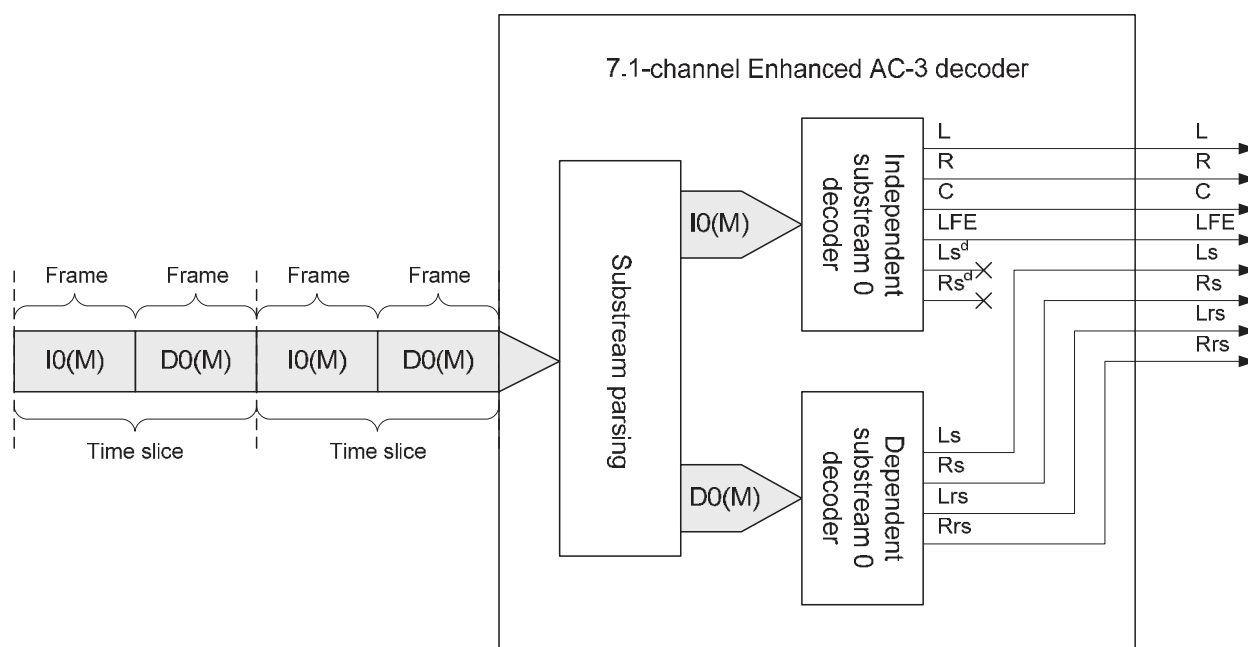


Figure 10.5 7.1-channel E-AC-3 decoder.

As mentioned in Section 10.1, multiple 7.1-channel configurations are supported by the E-AC-3 codec. Consequently, it is possible that the E-AC-3 decoder will receive as input a bitstream with a channel configuration (via chanmap and other metadata contained in the bitstream; see A/52 [1] E2.3.1.7, E2.3.1.8, and E3.8) that does not match the selected E-AC-3 decoder output configuration. In such a case, the E-AC-3 decoder simply decodes the 5.1-channel downmix contained in independent substream 0 and ignores dependent substream 0. If, under different circumstances, the channel configuration of the bitstream matches that of the selected E-AC-3 decoder output configuration, both independent substream 0 and dependent substream 0 are decoded.

10.4.2 5.1/2.0 Decode Process

When an E-AC-3 decoder receives a stream encoded with 7.1 channels of audio content but is capable of, or is configured to output, only 5.1 (or 2.0) channels of PCM audio, the E-AC-3 decoder will decode only independent substream 0. The dependent substream is ignored. All original sound

elements of the 7.1-channel mix are represented in the decoded 5.1 or 2.0-channel output because the independent substream contains a downmix of the 7.1-channel program, as described in Section 10.2.

Figure 10.6 shows an example of an E-AC-3 decoder decoding a 7.1-channel E-AC-3 stream and producing a 5.1 channel output. In this example, the decoder is only capable of decoding up to 5.1 channels from an E-AC-3 bitstream. The decoder reads the *strmtyp* and *substreamid* values of each frame, and processes only the frames that are from independent substream 0 (*strmtyp* = 0, *substreamid* = 0). The frames from dependent substream 0 (*strmtyp* = 1 *substreamid* = 0,) are ignored.

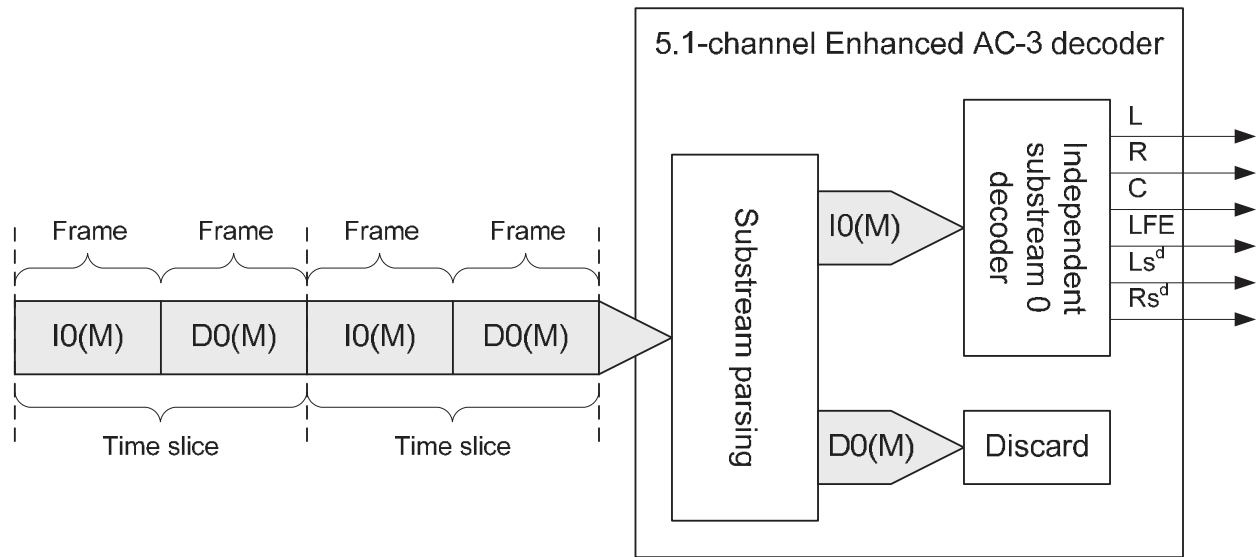


Figure 10.6 5.1-channel E-AC-3 decoder.

End of Document