

SMPTE STANDARD

for Television —

Mapping of AES3 Data into an MPEG-2 Transport Stream



Page 1 of 9 pages

Table of Contents

Page

1 Scope	1
2 Conformance Notation	1
3 Normative References	2
4 Introduction	2
5 SMPTE AES3 Elementary Streams	2
6 Packetized Elementary Stream	5
7 MPEG-2 Transport System	7
Annex A Bibliography (Informative)	8
Revision Notes	9

1 Scope

1.1 This standard specifies the method of transporting AES3 data in an MPEG-2 transport stream for television applications. For these television applications, the AES3 frame rate is 48 kHz.

1.2 Some applications may require linear PCM (pulse code modulated) digital audio in conjunction with compressed video specified in the MPEG-2 4:2:2 profile. The MPEG audio standard defines compressed audio, but does not define uncompressed audio for carriage in an MPEG-2 transport system. This standard augments the MPEG standards to address the requirement to carry AES3 streams, which may consist of linear PCM audio (at 48 ksample/s), or data carried within the AES3 format.

2 Conformance Notation

Normative text is text that describes elements of the design that are indispensable or contains the conformance language keywords: "shall", "should", or "may". Informative text is text that is potentially helpful to the user, but not indispensable, and can be removed, changed, or added editorially without affecting interoperability. Informative text does not contain any conformance keywords.

All text in this document is, by default, normative, except: the Introduction, any section explicitly labeled as "Informative" or individual paragraphs that start with "NOTE".

The keywords "shall" and "shall not" indicate requirements strictly to be followed in order to conform to the document and from which no deviation is permitted.

The keywords, "should" and "should not" indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.

The keywords "may" and "need not" indicate courses of action permissible within the limits of the document.

The keyword "reserved" indicates a provision that is not defined at this time, shall not be used, and may be defined in the future. The keyword "forbidden" indicates "reserved" and in addition indicates that the provision will never be defined in the future.

Unless otherwise specified the order of precedence of the types of normative information in this document shall be as follows. Normative prose shall be the authoritative definition. Tables shall be next, followed by formal languages, then figures, and then any other language forms.

3 Normative References

The following documents contain provisions which, through reference in this text, constitute provisions of this standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this standard are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

AES3-2003, Digital Audio Engineering — Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data

SMPTE 337M-2000, Television — Format for Non-PCM Audio and Data in an AES3 Serial Digital Audio Interface

SMPTE EG 32-1996, Emphasis of AES/EBU Audio in Television Systems and Preferred Audio Sampling Rate

ISO/IEC 13818-1:2000, Information Technology — Generic Coding of Moving Pictures and Associated Audio Information: Systems

4 Introduction

MPEG-2 transport streams convey one or more programs of coded data, and may be constructed from one or more elementary coded data streams, program streams, or other transport streams. This standard specifies the method to transport AES3 streams, which may consist of linear PCM digital audio, or data, within an MPEG-2 transport stream. The specifications are described in terms of a model which starts with AES3 data, constructs elementary streams (ES) from the AES3 data, then constructs packetized elementary streams (PES) from the elementary streams, and finally constructs MPEG-2 transport streams (MTS) from the packetized elementary streams. Although this model is used to describe the transport of AES3 streams in MPEG-2 transport streams, the model is not mandatory. MPEG-2 transport streams may be constructed by any method which results in a valid stream.

5 SMPTE AES3 Elementary Streams

5.1 SMPTE AES3 elementary streams shall consist of AES3 data words, which may be derived from AES3 subframes, together with validity, user, and channel status (V,U,C) bits and a framing (F) bit. The F bit

replaces the parity (P) bit normally carried in AES3 subframes, and is used to indicate the start of a block of AES3 channel status (C) and user (U) bit information (see paragraphs 5.5 and 5.6 and figure 3).

5.2 There may be 1, 2, 3, or 4 AES3 data streams conveyed in a single audio elementary stream and corresponding packetized elementary stream. Multiple packetized elementary streams may be used in applications requiring carriage of a larger number of AES3 streams. Where multiple packetized elementary streams are used to convey multiple AES3 streams, the ordering of AES3 subframes within a packetized elementary stream is maintained within the packetized elementary stream, while the ordering across different packetized elementary streams can be maintained through the value of channel_identification (see 6.6) in the elementary stream header.

5.3 The data word size shall be 16, 20, or 24 bits per word. All AES3 streams in one SMPTE AES3 elementary stream shall have the same word size.

5.4 The word rate shall be 48 kHz. This rate shall be locked to the 27-MHz transport clock. All channels in a SMPTE AES3 elementary stream shall have the same word frequency.

5.5 AES3 streams consist of a sequence of AES3 frames. A set of 192 sequential frames forms a block. Individual AES3 frames consist of a pair of AES3 subframes, designated the A subframe and the B subframe. AES3 subframes consist of data for a single data word, as well as additional data including a validity bit (V), a user data bit (U), a channel status bit (C), a parity bit (P), and four auxiliary sample bits (aux). AES3 subframes may include 24-bit data words as shown in figure 1, 20-bit data words as shown in figure 2, or 16-bit data words as shown in figure 3. AES3 subframes which carry 16-bit data words use a 20-bit word with the four least significant bits (LSBs) of the 20-bit word set to 0.

5.6 Validity, user, and channel status (V,U,C) bits shall be carried in the MPEG transport stream specified by this standard.

5.7 Framing information signaling the first AES3 subframe of a block is carried by a framing (F) bit that replaces the parity bit normally carried in the AES3 stream. The framing bit shall be set to 1 for the first A subframe of an AES3 block (indicated by the Z preamble in the AES3 stream) and to 0 for all other subframes (including all B subframes). Note that when multiple AES3 data streams are conveyed within one SMPTE AES3 elementary stream, each individual AES3 stream that is carried will have the F bit set to 1 for one A subframe out of every 192 A subframes; the F bit in all B subframes will always be set to 0. It is not necessary for all of the individual AES3 streams conveyed by the SMPTE AES3 data elementary stream have the same phasing of the 192 frame Blocks.

5.8 Elementary streams derived from AES3 data words and V,U,C,F bits shall be constructed as shown in figure 4. Elementary streams shall be composed of 302M word groups. Each 302M word group shall contain one AES3 data word together with the associated V,U,C and F bits for each subframe data channel carried in the elementary stream (i.e., one A subframe or one B subframe from each AES3 stream). The bits of the AES3 data words shall be the LSBs of the 302M data word; V shall be more significant, then U, then C, and the F bit shall be the MSB. The bits of the 302M data words shall be sent in the same order as in the AES3 sub-frame as shown in figures 1, 2, and 3 (i.e., LSB first). The 302M data words within each 302M data word group shall be in the order of their channel number (AES3 stream 1 subframe A, followed by AES3 stream 1 subframe B, followed by AES3 stream 2 subframe A, etc.). Successive 302M data word groups shall contain successive 302M data words from each of the channels carried in the elementary stream.