

SMPTE STANDARD

Format of Audio Metadata and Description of the Asynchronous Serial Bitstream Transport



Page 1 of 14 pages

Table of Contents	Page
Foreword	2
Intellectual Property	2
Introduction.....	2
1 Scope	3
2 Conformance Notation	3
3 Normative References	3
4 Definitions and Acronyms	4
5 Serial Metadata Stream Structure and Timing on a Serial Data Link	4
5.1 Structure on an Asynchronous Serial Line	4
5.2 Metadata Frame Timing Requirements	6
6 System Latency.....	7
7 Format of VANC Data Packets	7
7.1 SDID and Audio Channel Pair Number.....	7
8 Location of the Vertical Ancillary Data	8
9 Levels of Operation	9
9.1 Support Levels	9
9.2 Examples of Nomenclature	9
Annex A Bibliography (Informative)	10
Annex B Comments on the Preferred Location of VANC Packets (Informative)	11
Annex C Distinguishing Between Mapping Method A and Method B (Normative)	12
Annex D Asynchronous Serial Data Communication (Informative)	13
Annex E Assigning SDID Values (Normative)	14

List of Figures

Figure 1 – Structure of a Metadata Frame in the Asynchronous Serial Stream	5
Figure 2 – Asynchronous Serial Metadata Stream Timing	6

List of Tables

Table 1 – Association Between SDID Values and the First Audio Channel Pair of an Audio Program	8
--	---

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in its Standards Operations Manual.

SMPTE ST 2020-1 was prepared by Technology Committee 24TB.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Engineering Document. However, attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. SMPTE shall not be held responsible for identifying any or all such patent rights.

Introduction

This section is entirely informative and does not form an integral part of this document.

Audio data-rate reduction technologies can use metadata, which is multiplexed into the encoded audio bitstream, to describe the encoded audio and convey information that precisely controls downstream encoders and decoders.

Audio metadata is first created during program creation or mastering. It might need to be carried in the vertical ancillary data space of a digital television signal.

There is no specific provision in this standard for ensuring that the relative timing between the video and its embedded VANC data is correct. The only timing relationship that exists is created when the data is embedded in the VANC. Once that relationship is established, the deterministic nature of the serial digital interface ensures that the relationship is preserved

There are currently two methods of mapping the metadata into the vertical ancillary (VANC) data space. The following suite of SMPTE Standards, define the basic characteristics of the audio metadata, its serial transport used to convey the metadata between devices, and both of the mappings into the VANC space.

SMPTE ST 2020-1, Format of Audio Metadata and Description of the Asynchronous Serial Bitstream Transport

SMPTE ST 2020-2, Vertical Ancillary Data Mapping of Audio Metadata — Method A

SMPTE ST 2020-3, Vertical Ancillary Data Mapping of Audio Metadata — Method B

1 Scope

This Standard defines the basic structure and timing requirements of an asynchronous serial audio metadata stream with respect to an associated video signal. These requirements apply to the serial metadata streams either present on asynchronous serial data communications links between devices or present in the vertical ancillary (VANC) data packets transported by the video signal. The Standard also defines the Data ID (DID) and Secondary Data ID (SDID) values used to identify the VANC packets themselves and to associate the metadata in VANC packets with specific audio programs.

This Standard defines levels of operation for two methods of mapping the metadata into ancillary data packets.

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.

A conformant implementation according to this document is one that includes all mandatory provisions ("shall") and, if implemented, all recommended provisions ("should") as described. A conformant implementation need not implement optional provisions ("may") and need not implement them as described.

3 Normative References

The following standards 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.

ANSI TIA/EIA-485-A 1998 (R 2003), Electrical Characteristics of Generators and Receivers for Use in Balanced Digital Multipoint Systems

SMPTE ST 291-1:2011, Ancillary Data Packet and Space Formatting

SMPTE RP 168:2009, Definition of Vertical Interval Switching Point for Synchronous Video Switching

4 Definitions and Acronyms

4.1

Audio Channel

An audio signal representing the part of a program to be reproduced by one of the speakers, earpieces etc. of a sound reproduction system. A stereo program, for example, is composed of a Left channel audio signal and a Right channel audio signal.

4.2

Audio Channel Pair

Two audio signals representing part of or all of a program. The term may also be used to refer to the two signals carried by an AES3 digital audio interface.

4.3

Audio Program

A collection of one or more audio channels.

4.4

FPS

Frames per second.

4.5

Metadata Frame

A burst or frame of data that carries the metadata information related to a group of audio samples, which are temporally related to a video frame.

4.6

Serial Metadata Stream

The format and transport of a metadata frame as an asynchronous serial data signal on an ANSI/TIA/EIA-485-A physical interface.

4.7

VANC

Vertical Ancillary data space (see SMPTE ST 291-1).

5 Serial Metadata Stream Structure and Timing on a Serial Data Link

Audio metadata on a serial data link shall be transmitted as an intermittent, asynchronous serial data stream running at 115,200 baud, with 1 Start bit, 8 data bits, 1 Stop bit and no Parity bit, at TIA/EIA-485-A (formerly RS-485) electrical levels. The repetition rate of the Metadata frames shall match the repetition rate of the associated Interlaced or Progressive Video frames for video frame rates less than or equal to 30 frames per second and the repetition rate of pairs of Progressive video frames for rates greater than 30 frames per second as shown in Figure 2. The transmission of the first and second metadata subframes should be timed to avoid the areas around the vertical interval switch points of the associated video signal to avoid corrupting the metadata stream when it is switched synchronously with the video signal. The duration and timing of these intervals around the switch points allows some tolerance in the relative timing of the video and metadata streams. This is detailed in Section 5.2.

5.1 Structure on an Asynchronous Serial Link

Each metadata burst or frame shall carry the audio metadata information related to a specific group of audio samples, which are temporally related to a specific video frame or frames. Metadata frames shall be subdivided into a first and second subframe, each of which shall be prefixed by a specific sync sequence. The number of bytes in each subframe depends on the number and configuration of the associated rate reduced audio programs.

Figure 1 shows the general structure of the serial metadata stream. The duration of the subframes depends on the number and format of the associated audio programs, but the repetition period of the complete metadata frame shall be equal to the frame period of the associated video signal operating at or below 30 frames per second, or to a pair of frame periods of a progressively scanned video signal operating at greater than 30 frames per second. The top part of Figure 1 shows the structure of the serial metadata stream when the duration of Metadata Subframe 1 is less than one video field (or one Progressive frame). The bottom part of Figure 1 shows the structure when the duration of Metadata Subframe 1 is greater than one video field (or one Progressive frame).

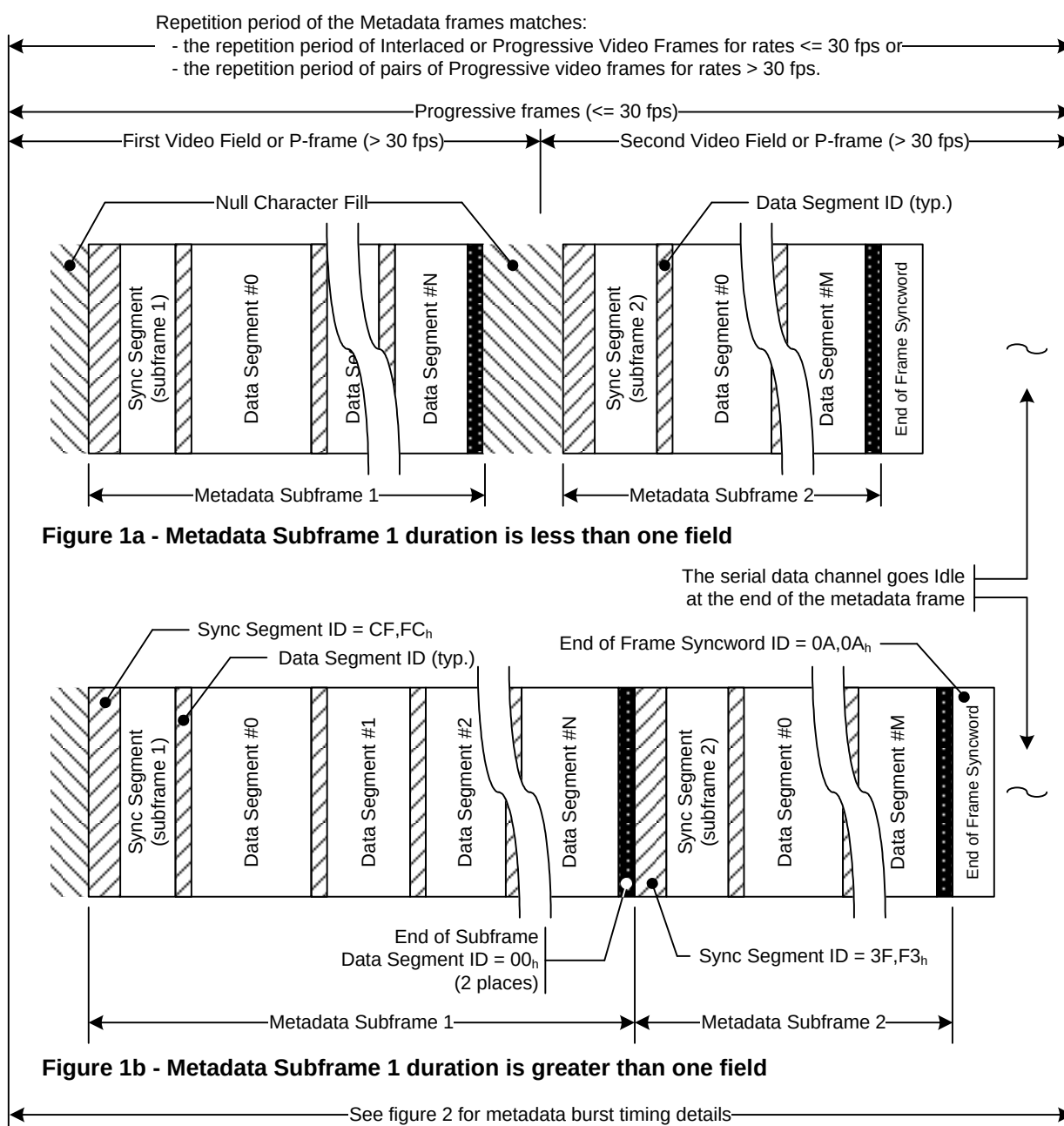


Figure 1 – Structure of a Metadata Frame in the Asynchronous Serial Stream