SMPTE ST 331:2011
Revision of
SMPTE 331M-2004

Element and Metadata Definitions for the SDTI-CP

Table of Contents

Approved
May 12, 2011

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 Part XIII of its Administrative Practices.

SMPTE ST 331 was prepared by Technology Committee 30MR.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Standard. 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.

1 Scope

This standard specifies the formats of the elements and metadata used by the SDTI content package format standard (SDTI-CP).

This standard defines element and metadata formats where they are simply specified or where a publicly available reference is available. It is not intended that this document provide detailed specifications for complex formats which may have broader application.

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.

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

Note: All references in this document to other SMPTE documents use the current numbering style (e.g. SMPTE ST 12-1:2008) although, during a transitional phase, the document as published (printed or PDF) may bear an older designation (such as SMPTE 12M-2-2008). Documents with the same root number (e.g. 12-1) and publication year (e.g. 2008) are functionally identical.

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 editions of the standards indicated below.

AES3-2009, AES Standard for Digital Audio Engineering — Serial Transmission Format for Two-Channel Linearly Represented Digital Audio Data

SMPTE ST 12-1:2008, Television — Time and Control Code

SMPTE ST 291:2010, Ancillary Data Packet and Space Formatting

SMPTE ST 309:1999, Television — Transmission of Date and Time-Zone Information in Binary Groups of Time and Control Code

SMPTE ST 312:2001, Television — Splice Points for MPEG-2 Transport Streams

SMPTE ST 326:2000, Television — SDTI Content Package Format (SDTI-CP)

SMPTE ST 328:2000, Television — MPEG-2 Video Elementary Stream Editing Information

SMPTE ST 330:2004, Television — Unique Material Identifier (UMID)

SMPTE ST 336:2007, Data Encoding Protocol Using Key-Length-Value

SMPTE ST 337:2008, Format for Non-PCM Audio and Data in an AES3 Serial Digital Interface

SMPTE RP 186:2008, Video Index Information Coding for 525- and 625-Line Television Systems

ISO/IEC 13818-2:2000, Information Technology — Generic Coding of Moving Pictures and Associated Audio Information: Video

4 Type Value of Element and Metadata

Each type of element and metadata listed below includes an element or metadata type value. This type value is not related to the data type values specified in SMPTE ST 305. The SDTI data type defines an item type, whereas this standard specifies element and metadata types.

The ranges of element and metadata type values for each item shall be as follows:

- Picture item: Element type range = 01_h to $0F_h$ inclusive.
- Audio item: Element type range = 10_h to $1F_h$ inclusive.
- Auxiliary item: Element type range = 20_h to 77_h inclusive.
- System item: Element type range = 78_h to $7F_h$ inclusive.
- System item: Metadata type range = 80_h to FF_h inclusive.

A type value of 00_h is not a valid value for either elements or metadata.

Elements types defined for use in the auxiliary item may be used in picture or audio items where so stated.

5 System Elements

5.1 Control Code Element

Type value = 78_h

Reserved but not yet defined.

6 Picture Elements

6.1 MPEG-2 Picture Element

Type value = 01_h

The MPEG-2 picture element shall be defined by the MPEG-2 video elementary stream (V-ES) of any profile or level according to ISO/IEC 13818-2.

The MPEG-2 picture element shall comprise the MPEG V-ES for one video frame together with all the MPEG-2 header information (including extensions) required to support the independent decoding of each picture.

The key start codes for MPEG-2 video elementary streams are:

- Sequence header (S): 00, 00, 01, B3;
- GOP header (G): 00, 00, 01, B8;
- Picture header (F): 00, 00, 01, 00;
- Sequence end code (E): 00, 00, 01, B7.

An example V-ES bitstream is shown in Figure 1. The MPEG-2 V-ES bit stream is simply formatted into a data block as indicated in the figure. No other data is required.

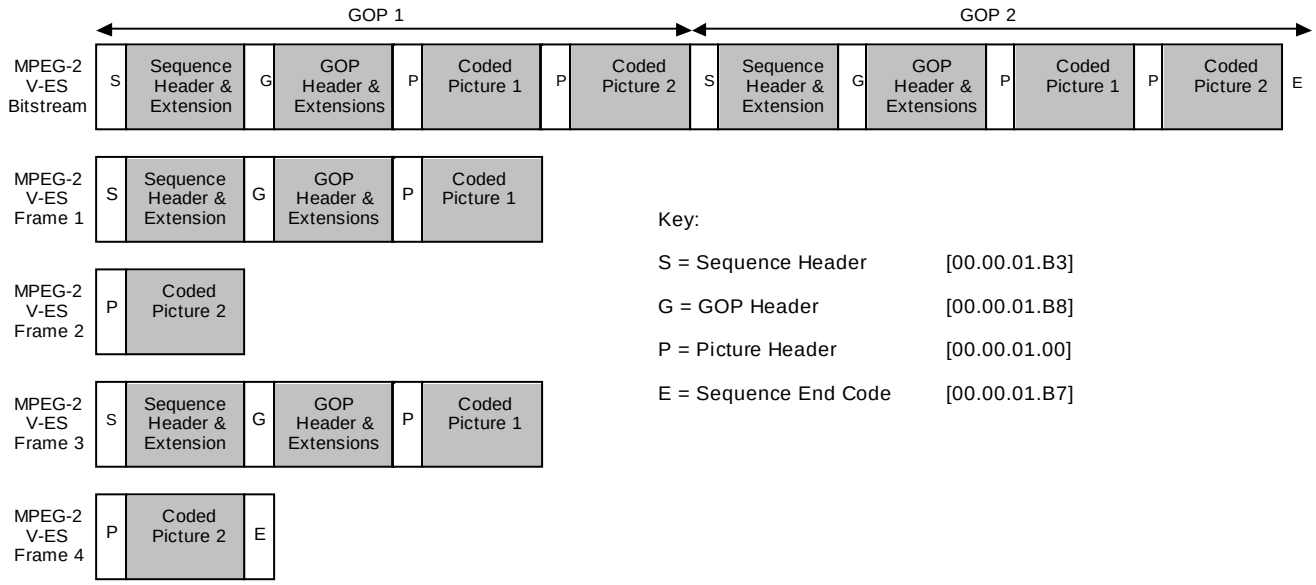

Figure 1 – Example formatting of a V-ES into MPEG-2 element frames

It is recommended that the MPEG-2 picture element complies with SMPTE ST 328 (MPEG-2 elementary stream editing information). The following informative list of points summarizes the provisions of that standard for the repetition of MPEG-2 GOP and sequence header information.

- If the picture to be formatted is not an I-picture, then the data from the picture header code up to, but not including, either the next GOP or picture header is formatted into a block;

- If the picture to be formatted is an I-picture, then the data from the sequence, GOP and picture headers up to, but not including, either the next GOP or picture header is formatted into a block;

- It is recommended that the sequence header information be repeated at each I-picture with the information placed immediately prior to the GOP header information. Thus information about the sequence is readily available following any editing process. If sequence header information were not repeated so frequently, then edit processes may easily remove this information making downstream processing more difficult or even impossible;

- A sequence end code shall be retained with the end of the last picture in the sequence. After editing, a new sequence end code shall be added to the end of a sequence if it does not already exist. The sequence of pictures is per the MPEG-2 picture bit stream including any discontinuity which may result from the use of B-pictures. Any picture stream timing metadata in the system Item shall reflect the decoded picture display sequence and thus the timing may appear discontinuous in the event of B-frames.

The byte alignment of both MPEG-2 and SDTI-CP is identical and this alignment is maintained in this standard. However, it should be noted that the bit stream orders of MPEG-2 and SDI differ in that the MPEG-2 bitstream is MSB first, whereas the SDI bit stream is LSB first.

Note: SMPTE RP 204 provides templates for the use of this element in SMPTE ST 326 together with a definition of a SMPTE label to identify the element complexity for interchange.

7 Audio Elements

7.1 8-Channel AES3 Element

Type value = 10_h

The data format of each channel of the 8-channel AES3 element is defined by the AES3 interface specification. Although the AES3 specification is limited to 2 channels, the 8-channel AES3 element is able to carry up to eight individual channels of AES3 data transparently. The I/O to the element will typically use AES3 twin-channel interfaces. Each AES3 channel may contain either linear PCM audio or data according to the AES3 specification.

The data format for an 8-channel AES3 element is shown in Figure 2. The element data area shall contain AES3 audio or data samples for the period of the picture frame as close as possible.

Up to 8 channels of AES3 data shall be multiplexed on a word-by-word basis; i.e., the first word (W) of each channel (Ch) is multiplexed into the sequence:

W1 Ch1, W1 Ch2, W1 Ch3, W1 Ch4, W1 Ch5, W1 Ch6, W1 Ch7, W1 Ch8

W2 Ch1, W2 Ch2, W2 Ch3, W2 Ch4, W2 Ch5, W2 Ch6, W2 Ch7, W2 Ch8

etc.

The format of the bits in each word is defined in Figure 2.

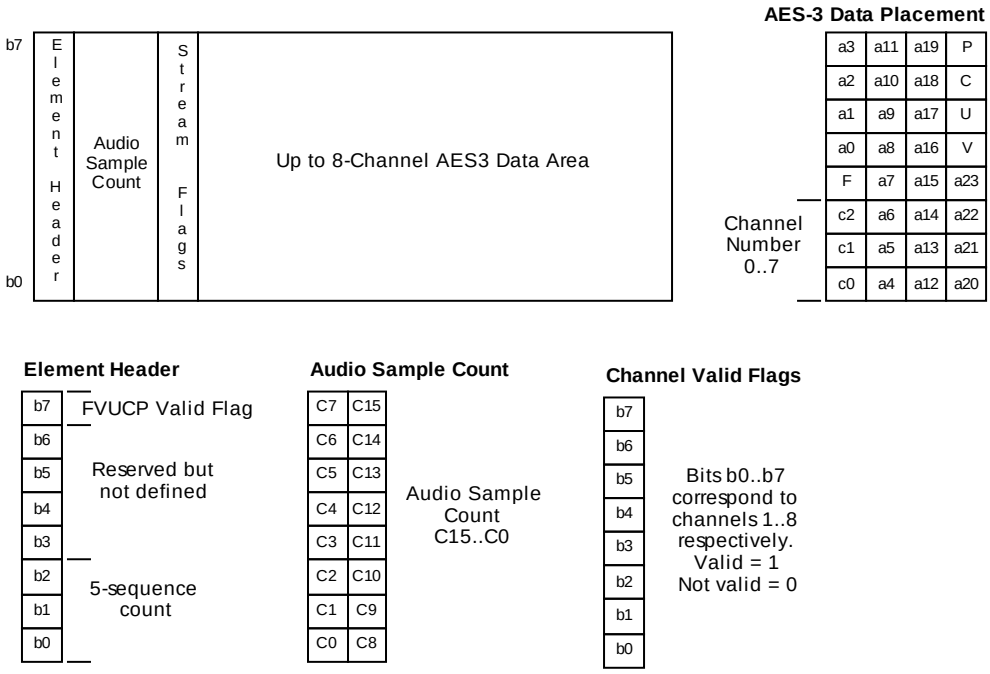

Figure 2 – Format of the 8-channel AES3 element

The channel number shall be defined by bits c2 to c0. These bits define 8 states where ‘0’ represents channel 1 and ‘7’ represents channel 8.

The F bit shall indicates the first AES3 sub-frame of an AES3 block. This bit shall be 1 if the word is the start of the AES3 block, else it shall be 0.

Note: Equipment exists with the previous definition of the F bit but because of the inability to apply this bit properly, there were at least two different interpretations. Equipment conforming to the above definition can be easily distinguished because the F bit is not static for every sample as was the case for the previous definition.

The 24 bits of the AES3 specification shall be directly mapped into bits a0 to a23. The V, U, C and P bits shall be directly mapped as shown in Figure 2.

For the AES3 element header:

- Bit b7 indicates if the FVUCP bits are active. A value of 0 indicates that the FVUCP bits are not used. A value of 1 indicates that the FVUCP bits are valid and useable.

- Bits b6 to b3 are not defined but reserved for future use.

- Bits b2 to b0 define a 5-sequence count. In a content package based on the 525/59.94 system, the count shall be a (modulo 5 + 1) count over the range 1 to 5. In a content package based on the 625/50 system, or any other system where the audio sample count is a consistent integer value over the content package period, the count shall be set to 0. All AES3 data channels within the same element shall have the same 5-sequence count number.

In the particular case of content packages based on 525/59.94 systems, the 5-sequence count defines one of the following sets of sample numbers per content package depending on whether it is frame or field based:

Sequence No.	30/1.001	60/1.001
1	1602	801
2	1601	801
3	1602	800
4	1601	801
5	1602	801

The audio sample count is a 16-bit count in the range 0 to 65535 and represents the number of samples in each channel. All channels within the element shall have the same sample count value.

The channel valid flag word has 8 bits, b0 to b7, which reflect the validity of the AES3 data in corresponding channels 1 to 8. A channel valid flag bit shall be set to 1 if the channel contains valid AES3 data else it shall be set to 0. The AES3 data area shall always be present for all 8 channels whether valid or not.

8 Auxiliary Elements

8.1 VBI Line Element

Type value = 20_h

The VBI line element carries one or more lines from the vertical blanking interval. The VBI line element has a header that identifies whether the source is interlaced or progressive and a number to identify the number of VBI lines carried.

Each VBI line is created from one line of the vertical blanking interval. Each line starts with a VBI information word followed by the 8-bit words from the whole of the VBI line.

Details of the VBI line element structure are shown in Figure 3.

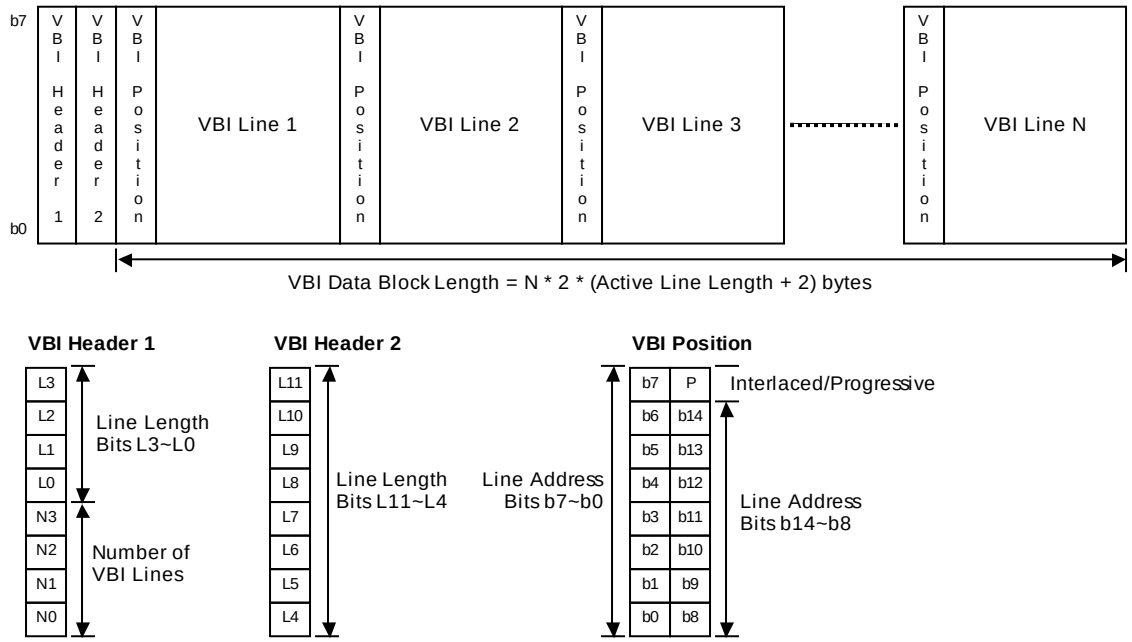

Figure 3 – Format of the VBI line element

The order of the VBI lines shall be as they are displayed on a viewing device.

For an interlaced scanned system, this shall be in the following order:

[VBI 1, 1st field], [VBI 2, 2nd field],

[VBI 3, 1st field], [VBI 4, 2nd field],

[VBI 5, 1st field], [VBI 6, 2nd field].

For a progressive scanned system, this shall be in the following order:

VBI 1, VBI 2, VBI 3, VBI 4, VBI 5, VBI 6.

In the VBI header words:

- Bits N3 to N0 of the first word shall define the number of VBI lines. The allowable values shall be in the range 0 to 6.

- Bits L3 to L0 of the first word together with bits L8 to L11 of the second word form a 12 bit count value which identify the length of the VBI lines. All VBI lines in one element shall have the same length.

In the VBI position word:

- Bits b14 to b8 of the second word and bits b7 to b0 of the first word form a line number range of 0 to 32767. The line address number shall represent an absolute line number for both Interlaced and Progressive line numbering systems.

- Bit b7 of the second word (P) is set to ‘0’ for Interlaced scan and ‘1’ for progressive scan.

A line address value of "0" means that no line number has been defined. Any line address number outside the vertical interval period for the picture scanning system is invalid and may cause unspecified effects in receiving equipment.

8.2 Ancillary Data Packet Element

Type value = 21_h

Ancillary data packets are defined by SMPTE ST 291. This format carries only the 8 LSBs of each word of the ANC data packets and removes the ADF word sequence. Where there is more than one ANC packet, they shall be packed in sequence with no padding words or gaps between the packets. Reformatting to the full 10-bit word resolution together with the addition of the ADF sequence to the head of each ANC packet is the responsibility of the output formatting device.

Ancillary data packets carry data identification codes (DID and SDID) which identify the type of payload. The values for these codes are defined in the appropriate SMPTE standards and recommended practices.

The format of a type 2 ancillary data packet is shown in Figure 4. The format may carry type 1 ancillary data by replacing the secondary DID word with the data block number.

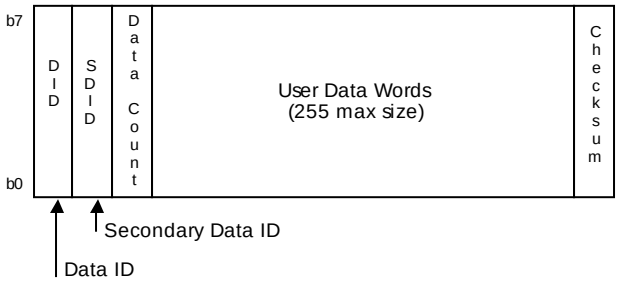

Figure 4 – Format of the ancillary data packet element (Type 2 shown)

8.3 General Data Element

Type value = 22_h

The general format shown in Figure 5 is used to carry all free-form data types which do not have a separate auxiliary item element type value. These data types, including those of an IT nature (Word processing files, Hypertext etc), may require format identification through associated metadata in the system item.

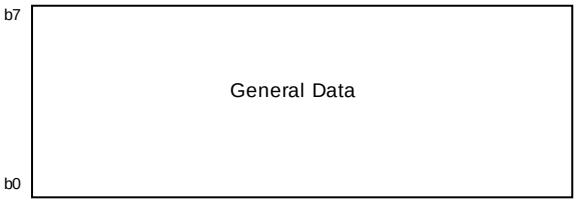

Figure 5 – Format of the general data element

The auxiliary item may contain several general data elements with the same type identifier. These are distinguished with the element number which shall match the element type and number used in any associated metadata packet.

8.4 BWF Element

Type value = 40_h

Reserved for the broadcast wave format (BWF), but not defined.

8.5 JFIF Element

Type value = 41_h

Reserved for the JPEG file interchange format (JFIF) element, but not defined.

8.6 TIFF Element

Type value = 42_h

Reserved for the tagged image file format (TIFF), but not defined.

9 System Item Metadata Definitions

The following metadata definitions may be used in the system item to describe, as appropriate, either the entire content package, or any element or combination of elements in the picture, audio, or auxiliary items.

Note: Metadata definitions have both a local identifier defined by the type value and a global identifier which is the key in the SMPTE Data Element Dictionary (SMPTE RP 210). Both identifiers are referencing the same metadata specification defined in this document. The reason for the short type value used in the SDTI content package is for ease of parsing the data at the high speeds used by the SDTI transport. There is also a gain in packing density and hence simplified storage requirements on high-speed silicon. Most metadata items specified in this standard can be expanded to the full K-L-V construct using the key from the metadata dictionary. This fully expanded K-L-V construct can then be used as a basis for the common interchange of metadata items between different applications.

The first 8-byte group of the key is the SMPTE metadata dictionary designator value and is set to the value:

06.0E.2B.34.01.01.01.vv_h

where vv_h is the version of the SMPTE metadata dictionary at the time of registration of the item.

The second 8-byte group of the key is the metadata dictionary tag value and this is the value indicated in the metadata definitions below. Some metadata items defined below have specialized and complex data structures with no entry in the metadata dictionary.

9.1 Metadata Link Item

Local type value = 80_h

Tag value = 01.07.01.03.00.00.00.00_h

This special metadata item shall be used to link subsequent metadata items to their respective elements. The format differs from the general metadata construct and consists simply of 3 bytes which shall be in the sequence:

metadata type, element type, element number.

In any picture, audio or auxiliary metadata set, a metadata link item shall be present immediately following any nonzero metadata count value. All metadata following a metadata link item shall refer to the defined element until the occurrence of the next metadata link item.

9.2 SMPTE ST 12-1 Time Code Metadata

Type value = 81_h

Tag value = 07.02.01.01.01.04.00.00

This metadata consists of a 16 byte field with the first 8 bytes coded with the data from the SMPTE ST 12-1 time code specification and is shown in Figure 6. The last 8 bytes of the metadata format shall be null filled.

Note that the metadata is organized as LSB first to comply with SDI convention.

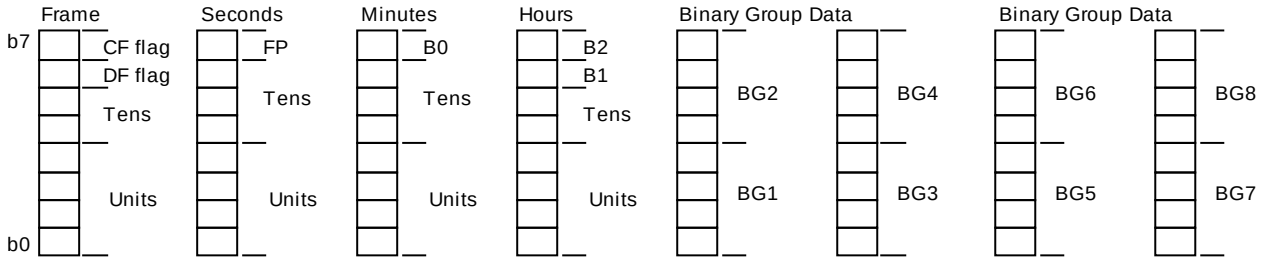

Figure 6 – Illustration of the time code data format

The order of transmission is bit b0 of the leftmost word first finishing with bit b7 of the rightmost word.

A list of the abbreviated terms in Figure 6 and their full names follows:

- CF flag: Color frame flag;
- DF flag: Drop frame flag;
- FP: Field phase (NTSC), Binary group flag 0 (PAL);
- B0: Binary group 0 (NTSC), Binary group 2 (PAL);
- B1: Binary group 1 (NTSC and PAL);
- B2: Binary group 2 (NTSC), Field phase (PAL).

9.3 SMPTE ST 309 Date-Time Stamp Metadata

Type value = 82_h

Tag value = 07.02.01.01.01.03.00.00

This metadata consists of a 16-byte field with the first 8 bytes coded according to the SMPTE ST 309 specification with a time, time zone, and date fields. The last 8 bytes of the metadata format shall be null filled.

The mapping of data from SMPTE ST 309 to the SDTI-CP metadata format is identical to that described in Section 9.2 (SMPTE ST 12-1 time code).

9.4 SMPTE UMID Metadata

Type value = 83_h

Tag value = 01.01.0t.mi.00.00.00.00

(where: t = essence type, m = material number creation method and i = instance number creation method).

The UMID is defined in SMPTE ST 330. It may be a basic UMID with a total length of 32 bytes or an extended UMID with a total length of 64 bytes. Both UMID types are defined as an integrated K-L-V construct with the last 4 bytes of the key discarded.

9.5 MPEG-2 Picture Editing Metadata

Type value = 84_h

Tag value: Not available.

Figure 7 illustrates the data structure of the MPEG-2 picture editing metadata.

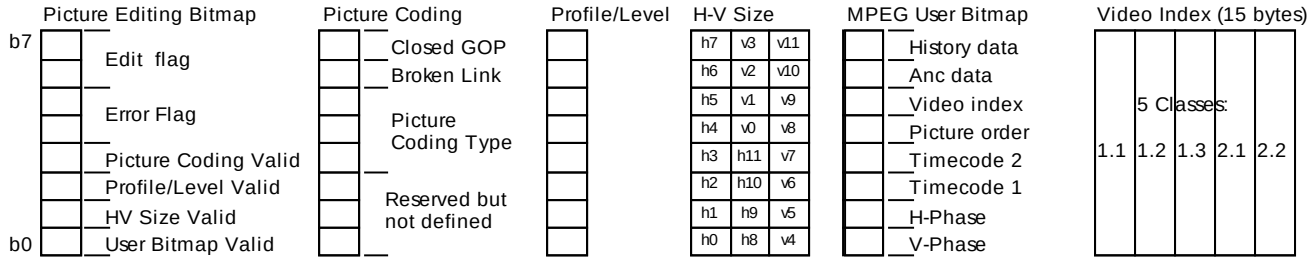

Figure 7 – Format of the MPEG-2 picture editing metadata

The transmission order is LSB first to comply with the SDTI specification.

The format is now defined in pseudo-code representation based on byte transmission order. Values in square brackets indicate the size of the data field in bytes. Note that the data within a byte is ordered in relation to the diagram and does not imply the order of bit stream transmission.

```
MPEG-2_Picture_Editing_Metadata()
{
    picture_edit_bitmap [1]
    picture_coding_parameters [6]
    video_index [15]
    extension_data [var]
}

picture_edit_bitmap()
{
    edit_flag                2 bits
    error_flag               2 bits
    picture_coding_valid     1 bit
    profile_level_valid      1 bit
    HV_size_valid            1 bit
    user_bitmap_valid        1 bit
}

edit_flag (bits b7 b6)
00 : No edit
01 : Pre-picture edit
        (the previous element is no longer related to the current element)
10 : Post-picture edit
        (the following element is no longer related to the current element)
11 : Single frame picture
        (neither the previous nor the following elements are related to the current element)

error_flag (bits b5 b4)
00 : error status not known
01 : concealed error
10 : uncorrected error
11 : no error

if(picture_coding_valid == 1)
{
    picture_coding_data()
    {
        closed_gop           1 bit (as MPEG)
        broken_link          1 bit (as MPEG)
```

```
            picture_coding_type             3 bits (as MPEG, MSB in bit b5)
            reserved but not defined        3 bits
        }
    }
    else
        picture_coding_data [1]: non valid data

    if(profile_level_valid == 1)
    {
        profile_level           8 bits (as MPEG, MSB in bit b7)
    }
    else
        profile_level [1]: nonvalid data

    if(HV_size_valid == 1)
    {
        HV_size()
        {
            horizontal_size         12 bits (bits h11 to h0 as defined in Figure 7)
            vertical_size           12 bits (bits v11 to v0 as defined in Figure 7)
        }
    }
    else
        HV_size [3]: non valid data

    if(user_bitmap_valid == 1)
    {
        mpeg_user_data_bitmap()
        {
            History Data:     active = 1, not active = 0
            Ancillary Data:   active = 1, not active = 0
            Video Index:      active = 1, not active = 0
            Picture Order:    active = 1, not active = 0
            Timecode 2:       active = 1, not active = 0
            Timecode 1:       active = 1, not active = 0
            H-Phase:          active = 1, not active = 0
            V-Phase:          active = 1, not active = 0
        }
    }
    else
        mpeg_user_data_bitmap [1]: non valid data
    }

    NOTE – MPEG User Data is specified in SMPTE ST 328.

    video_index()
    {
        video_index_class_1.1 [3]
        video_index_class_1.2 [3]
        video_index_class_1.3 [3]
        video_index_class_2.1 [3]
        video_index_class_2.2 [3]
    }
```

Note: Video_index_class data is defined in SMPTE RP 186.

9.6 8-Channel AES3 Editing Metadata

Type value = 85_h

Tag value: Not available.

Figure 8 illustrates the data format of the 8-channel AES3 editing metadata. This metadata is designed to describe the 8-channel AES3 element (type 10_h). The channel status data shall represent the default channel status of the audio data streams within the 8-channel AES3 element.

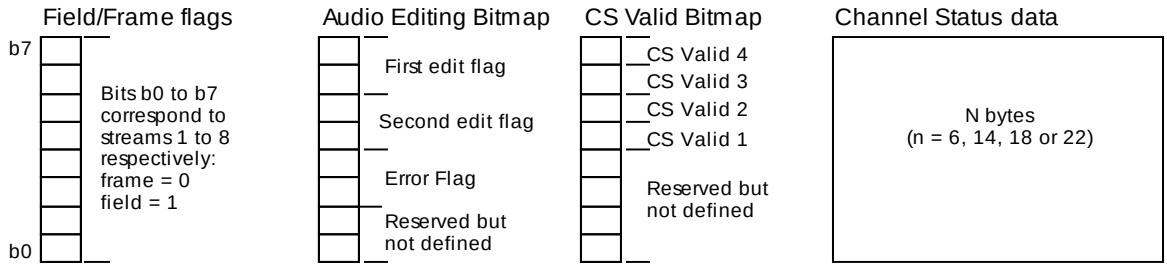

Figure 8 – Format of the 8-channel AES3 editing metadata

The transmission order is LSB first to comply with the SDTI specification.

The format is now defined in pseudo-code representation based on byte transmission order. Values in square brackets indicate the size of the data field in bytes. Note that the data within a byte is ordered in relation to the diagram and does not imply the order of bit stream transmission.

```
8-Channel_AES3_ Editing_Metadata()
{
    field_frame_flags [1]

    for(ch_number=0; ch_number<8; Ch_number++)
    {
        if(stream_valid_flag[ch_number] == 1)
        {
            first_edit_flag                 2 bits
            second_edit_flag                2 bits
            error   _flag                   2 bits
            reserved but not defined        2 bits
        }
        else
            null_data                       8 bits
    }

    for(ch_number=0; ch_number<8; Ch_number++)
    {
        if(stream_valid_flag[ch_number] == 1)
        {
            cs_valid_4                      1 bit
            cs_valid_3                      1 bit
            cs_valid_2                      1 bit
            cs_valid_1                      1 bit
            reserved but not defined        4 bits
```

```
            if(cs_valid_1 == 1)
            {
                for(n=0; n<6; n++)
                    cs_byte(n) [1]
            }

            if(cs_valid_2 == 1)
            {
                for(n=6; n<14; n++)
                    cs_byte(n) [1]
            }

            if(cs_valid_3 == 1)
            {
                for(n=14; n<18; n++)
                    cs_byte(n) [1]
            }

            if(cs_valid_4 == 1)
            {
                for(n=18; n<22; n++)
            cs_byte(n) [1]

            }
        }
    }
}
```

Note: When a picture element has 2 fields, there are 2 audio edit positions defined by the positions of each picture field. The first and second edit flags apply to the respective first and second picture fields. When a picture is a single frame, only the first edit flag value is valid. Each bit in the field_frame_flags is used to identify whether the corresponding channel is edited on a field or frame basis.

first_edit_flag (bits b7 b6)
00 : No edit
01 : Pre-picture edit(the previous element is no longer related to the current element)
10 : Post-picture edit(the following element is no longer related to the current element)
11 : Single frame picture (neither the previous nor the following elements are related to the current element)

second_edit_flag (bits b5 b4)
00 : No edit
01 : Pre-picture edit(the previous element is no longer related to the current element)
10 : Post-picture edit(the following element is no longer related to the current element)
11 : Single frame picture (neither the previous nor the following elements are related to the current element)

error_flag (bits b3 b2)
00 : error status not known
01 : concealed error
10 : uncorrected error
11 : no error

Note: Stream_valid_flag[ch number] indicates the presence of an audio channel as defined by the stream valid flag bitmap in the 8-channel AES3 element.

9.7 Picture Bitstream Splicing Metadata

Type value = 86_h

Tag value = $05.02.01.02.01.01.00.00_h$

This metadata specification may be used to describe splicing information for MPEG video elementary streams with a GOP >1.

The format is defined as follows, following MPEG-2 notation:

```
Picture_Bitstream_Splicing_Metadata()
    {
        In_Point_Present [1]
        Out_Point_Present [1]
        Reserved [6]

        If(In_Point_Present)
        {
            In_Point_Splice_Type [4]
            Reserved [3]
            Closed_GOP [1]
        }

        If(Out_Point_Present)
        {
            Out_Point_Splice_Type [4]
            Reserved [4]
        }
    }
```

If the in_point_present is 1, the associated picture element shall commence with an in_point as defined by SMPTE ST 312 with the following exception:

- The last sentence of Section 5.3.2.1 does not apply. The effect of this exception is to allow an open GOP at an in_point. The closed_GOP flag shall be used to indicate whether a closed or open GOP exists at the in_point.

If the out_point_present is 1, it indicates that the associated picture element ends with an out_point as defined by SMPTE ST 312.

The in_point_splice_type is equal to the value of the splice_type defined in SMPTE ST 312 and indicates that the video elementary stream meets the constraints corresponding to the value defined in Sections 5.3.2.2 and 5.3.2.3 of SMPTE ST 312.

If closed_GOP is 1, it indicates that the in_point meets the constraints defined in SMPTE ST 312, Section 5.3.2.1.

The out_point_splice_type is equal to the value of the splice_type as defined in SMPTE ST 312, Section 5.2.2 as it pertains to the video elementary stream.

9.8 MPEG Decoder Buffer Delay Metadata

Type value = 87_h

Tag value = $07.02.03.01.03.01.00.00_h$

The decoder buffer delay metadata shall only be used to support the low latency transfer mode of MPEG video elementary streams. Its presence in low latency mode transfer provides an explicit definition of decoder buffer delay. If low latency mode is used without this metadata, the buffer delay is not defined and a decoder shall set its delay according to its own capabilities.

The decoder buffer delay metadata may be used to support low latency transfer mode for any picture, audio or auxiliary data element. The delay shall be specified as a 16-bit word representing the decoder buffer delay as a count of 90-kHz clock periods as follows:

- First byte: Bits b0 to b7 of the 16-bit count word, where bit b0 is the LSB of the first byte.
- Second byte: Bits b8 to b15 of the 16-bit count word, where b8 is the LSB of the second byte.

The 90-kHz clock shall be derived either from the 27-MHz SDTI clock by dividing it by 300, or from the 36-MHz SDTI clock by dividing it by 400.

The decoder buffer delay metadata value shall be used to define the period, in 90-kHz clock cycles, between writing the first word of an element into the decoder buffer and reading the first word of that element out of the decoder buffer. The decoder buffer delay value shall be defined as a value large enough to prevent decoder buffer underflow. The maximum decoder buffer requirement to prevent decoder buffer overflow is not specified.

9.9 KLV Metadata

Type value = 88_h

Tag value = $03.01.02.10.02.00.00.00_h$

The KLV metadata format shown in Figure 9 is defined for the carriage of any metadata that is KLV coded according to SMPTE ST 336. This metadata item shall be used for the coding of the whole of a single KLV data item or a single KLV data group.

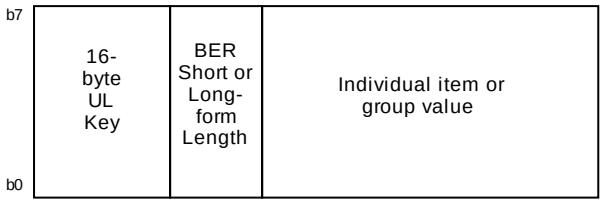

Figure 9 – KLV metadata format

The key shall be a 16-byte SMPTE UL as defined in SMPTE ST 336.

The length field may be long or short form BER coded as defined by SMPTE ST 336.

The value is coded as an item or a group according to the UL designator of the key as defined in SMPTE ST 336.

The length of the KLV length field shall be constrained to be within the capability of the 4-byte element word count as defined in SMPTE ST 326.

9.10 AES3 Non-Audio Metadata

Type value = 89_h

Tag value: Not available

This metadata specification may be used to describe individual channels of AES3 non-audio data in the 8-channel AES3 element (type 10h).

Note: If the 8-channel AES3 element carries 4 channels of AES3 non-audio data, then there will be 4 AES3 non-audio metadata items, one to describe each channel.

The AES3 non-audio metadata format shown in Figure 10 identifies the payload of any AES3 non-audio data. It shall not be used to describe AES3 audio data.

Individual fields indicated in Figure 10 are defined by SMPTE ST 337. The values of certain fields may be specified by related SMPTE standards.

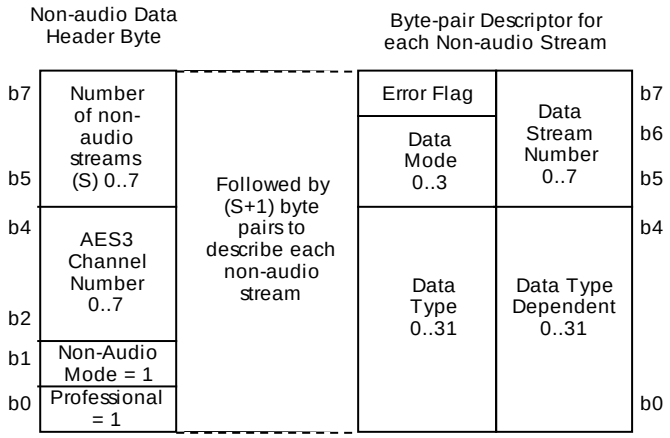

Figure 10 – Format of AES3 non-audio metadata

The AES3 non-audio metadata shall start with a single header byte that defines the following:

- Bit b0 shall be set to ‘1’ to indicate professional use of the AES3 data.
- Bit b1 shall be set to ‘1’ to indicate non-audio use of the AES3 data.
- Bits b2 to b4 shall define the channel number in the 8-channel AES3 element.
- Bits b5 to b7 shall define the number of non-audio data streams are contained within this AES3 channel.

The header byte shall be followed by a sequence of one to eight byte pairs where each byte pair uniquely describes a non-audio stream. Each byte pair contains a copy of the data fields defined in the burst_info of SMPTE ST 337.

- Bits b0 to b4 of the first byte shall be set to indicate the “data_type” value as specified by SMPTE ST 337.
- Bits b5 and b6 of the first byte shall be set to indicate the “data_mode” value as specified by SMPTE ST 337.
- Bit 7 of the first byte shall be set to ‘1’ to indicate that the “error_flag” as specified by SMPTE ST 337 has been set to ‘1’ within this stream of this channel of the AES3 Element.
- Bits b0 to b4 of the second byte shall be set to indicate the “data_type_dependent” value as specified by SMPTE ST 337.
- Bits b5 to b7 of the second byte shall be set to indicate the “data_stream_number” of this stream as specified by SMPTE ST 337.

Annex A Bibliography (Informative)

SMPTE ST 305:2005, Television — Serial Data Transport Interface (SDTI)

SMPTE ST 363.2-2002, Television — Declarative Data Essence, Content Level 1 [Appendix D for JFIF]

SMPTE RP 204:2000, SDTI-CP MPEG Decoder Templates

SMPTE RP 210, Data Element Dictionary

ISO/IEC 10918-1:1994, Information Technology — Digital Compression and Coding of Continuous-Tone Still Images: Requirements and Guidelines [for JPEG]

ISO/IEC 10918-3:1997, Information Technology — Digital Compression and Coding of Continuous-Tone Still Images: Extensions [for JPEG]

ITU-R BR.1352-3:2007, File Format for the Exchange of Audio Programme Materials with Metadata on Information Technology Media, Annex 1 Specification of the Broadcast Wave Format

www.libtiff.org/document.html [for TIFF]