SMPTE ST 381-3:2017
Revision of
ST 381-3:2013

SMPTE STANDARD

Material Exchange Format—Mapping AVC Streams into the MXF Generic Container

Table of Contents

Approved August 30, 2017

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 Operations Manual.

SMPTE ST 381-3 was prepared by Technology Committee 31FS.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential for 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.

1 Scope

This standard specifies the mapping of all AVC coding data as defined in ISO/IEC 14496-10 | Rec. ITU-T H.264 into the Material Exchange Format Generic Container (MXF-GC) based on the MXF MPEG mapping standard (SMPTE ST 381-2).

This standard does not apply to Annexes G (SVC) or H (MVC) of the AVC standard. The mapping of SVC and MVC might be covered by other SMPTE documents. This standard also does not support the AVC file format (ISO/IEC 14496-15).

This document specifies the carriage of parameter sets in an AVC essence stream “In-band”.

The MXF specification is written in several parts. This is an MXF mapping specification that defines header metadata sets and values, essence container elements and index table applications.

In order to achieve interoperability within any given Operational Pattern, restrictions may be placed on the way in which this Generic Container type can be implemented. The reader is advised to carefully study the appropriate Operational Pattern document before implementation.

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 379-1:2009) although, during a transitional phase, the document as published (printed or PDF) may bear an older designation (such as SMPTE 379-1-2009). Documents with the same root number (e.g. 379-1) and publication year (e.g. 2009) 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 edition of the standards indicated below.

SMPTE ST 377-1:2011, Material Exchange Format (MXF) – File Format Specification

SMPTE ST 379-1:2009, Material Exchange Format (MXF) — MXF Generic Container

SMPTE ST 379-2:2010, Material Exchange Format (MXF) – MXF Constrained Generic Container

SMPTE ST 381-2:2011, Material Exchange Format (MXF) – Mapping MPEG Streams into the MXF Constrained Generic Container

ISO/IEC 13818-1:2015, Information technology – Generic coding of moving pictures and associated audio information: Systems

ISO/IEC 14496-10:2014 | Rec. ITU-T H.264 (2016), Information technology — Coding of audio-visual objects – Part 10: Advanced Video Coding or ITU-T Recommendation H.264 — Advanced Video Coding for Generic Audio-Visual Services

4 Definition of Acronyms, Terms and Data Types

4.1 The general glossary of acronyms, terms and data types used in the MXF specification is given in SMPTE ST 377-1, SMPTE ST 379-1 and SMPTE ST 379-2. Definitions of terms, abbreviations and symbols relating to AVC are given in ISO/IEC 14496-10 | Rec. ITU-T H.264.

4.1.1
AVC
Advanced Video Coding as documented in ISO/IEC 14496-10 I Rec. ITU-T H.264 (also termed MPEG-4 Part 10)

4.1.2
NAL
Network Abstraction Layer

4.2 The following terms are not defined in the AVC specification, but are defined here as properties in the Descriptor and the Index Table Entry. The terms are applicable only to the MXF AVC wrapping defined in this standard.

4.2.1
Picture
a collective term for a frame or a field

4.2.2
I Picture
picture consisting only of I-Slices

4.2.3
non-IDR I Picture
I Picture other than IDR-Picture

4.2.4
P Picture
picture consisting of P-Slices, or P-Slices and I-Slices

4.2.5
B Picture
picture consisting of B-Slices, or B-Slices and P-Slices, or B-Slices, P-Slices and I-Slices

4.2.6
Br Picture
B Picture with nal_ref_idc nonzero. A B Picture that is available as a reference for another Picture

4.2.7
GOP
this term is not defined in the AVC specification, but is widely used in the industry, and is defined for the purposes of this standard: Group of Pictures starting with I Picture in coded order

4.2.8
Key Picture
the earliest preceding I Picture required for decoding the indexed Picture

4.2.9
Parameter sets
sequence parameter sets and picture parameter sets

4.2.10
In-band
parameter sets are carried in an AVC essence stream

5 Technical Introduction (Informative)

5.1 AVC Coding Summary

The AVC (Advanced Video Coding) standard defined in the ISO/IEC 14496-10 | Rec. ITU-T H.264 has been developed as an advanced high performance compression technology. In the AVC coding streams, the bit stream can be in one of two formats: the NAL unit stream or the byte stream.

The NAL unit stream format consists of a sequence of NAL unit syntax structures.

The byte stream format can be constructed from the NAL unit stream by prefixing each NAL unit with a start code prefix and zero or more zero-valued bytes to form a stream of bytes.

5.2 Requirements of Mapping AVC Streams into MXF

The specification on the MXF AVC mapping was developed to meet the following requirements:

- The mapping needs to support the NAL unit stream format and the byte stream format.
- AVC streams need to be wrapped using the MPEG Picture Element Key as defined in SMPTE ST 381-2.
- The Picture Element can contain all variants of AVC coding streams using frame wrapping, clip wrapping, or custom wrapping.
- The mapping needs to support carriage of sequence parameter sets and picture parameter sets “In -band“.
- Full details of the precise AVC stream need to be specified in the Essence Descriptor, and AVC coding variants need to be specified in the Picture Essence Coding Label.
- A Sub-Descriptor needs to be specified to give information on AVC streams.

This document meets the above requirements for the MXF mapping of the AVC streams.

6 Mapping the AVC Streams to the MXF Generic Container

AVC streams shall be mapped using the MXF Generic Container in the Frame-based wrapping, Clip-based wrapping or the Custom wrapping as defined in SMPTE ST 379-1 and SMPTE ST 379-2.

7 Key-Length-Value Coding

7.1 Picture Element Key

AVC streams shall be wrapped using the MPEG Picture Element Key as defined in SMPTE ST 381-2. The values of the first 12 bytes of the Essence Element Key are defined in SMPTE ST 379-1 and SMPTE ST 379-2. The values of the last four bytes of the Picture Element Key are given in Table 1.

Table 1 – Key Value for the MPEG Picture Element (Informative)

Byte No.	Description	Value (hex)	Meaning
1-12	Specified by the MXF Generic Container Specification, SMPTE ST 379-1 and SMPTE ST 379-2		
13	Item Type Identifier	15_h	Generic Container Picture Item (as defined in SMPTE ST 379-1 and SMPTE ST 379-2)
14	Essence Element Count	kk_h	Count of Picture Elements in this Picture Item
15	Essence Element Type	05_h 06_h 07_h	Frame Wrapped Picture Element Clip Wrapped Picture Element Custom Wrapped Picture Element
16	Essence Element Number	nn_h	The Number (used as an Index) of this Picture Element in this Picture Item

7.2 Picture Element Length

The length field of the KLV coded Element shall be 4 bytes BER long-form encoded (i.e. 83h.xx.yy.zz) for Frame wrapping. The length field of the KLV coded Element shall be 8 bytes BER long-form encoded (i.e., 87h.aa.bb.cc.dd.ee.ff.gg) for Clip wrapping. For Custom wrapping, the length field shall be constant for all Edit Units of the essence container and shall be either 4 bytes or 8 bytes as appropriate.

7.3 Picture Element Value

The Picture Element Values shall be the AVC NAL unit stream or the AVC byte stream. The bit streams carried in the Value field shall contain complete NAL units including their relevant parameter sets, other Supplemental Enhancement Information (SEI) and padding zeroes.

Note: Stream format is defined in the Byte 14 of the Essence Container Label described in Section 8.1.

If a field-encoded stream is frame-wrapped, the Value field shall contain two fields, with a frame comprising a pair of fields, and byte 15 of the Picture Element Key shall be 05h (frame-wrapped). The Edit Rate of the Track that links such an Essence Container Element shall equal the frame rate.

If a field-encoded stream is field-wrapped, the Value field shall contain a single field, and byte 15 of the Picture Element Key shall be 05h (frame-wrapped). The Edit Rate of the Track that links such an Essence Container Element shall equal the field rate.

The frame-wrapping or the field-wrapping shall be signaled by the Essence Container Label (see Table 2).

8 SMPTE Label Values

8.1 Essence Container Label

This Essence Container Label is the UL value carried in the Essence Containers Properties of the Partition Packs, Preface Set and File Descriptor.

The values of the Essence Container Label for the AVC coding shall be as shown in Table 2.

Table 2 – Specification of the AVC Essence Container Label

Byte No.	Description	Value (hex)	Meaning
1-12	Specified by the MXF Generic Container Specification, SMPTE ST 379-1 and SMPTE ST 379-2		
13	Essence Container Kind	02_h	MXF Generic Container (as defined in SMPTE ST 379-1 and SMPTE ST 379-2)
14	Mapping Kind	$0F_h$ 10_h	AVC NAL unit stream AVC byte stream
15	Locally defined	xx_h	ISO/IEC 13818-1 stream id bits 6…0 The default value is 60h.
16	Locally defined	yy_h	00h: Not used 01h: Frame Wrapping 02h: Clip Wrapping 03h Custom: Stripe Wrapping 04h: Reserved 06h Custom: Splice 07h Custom: Closed GOP 08h Custom: Slave 09h: Frame (Field) Wrapping 0Ah - 7Eh: Reserved

Byte 14 identifies the container as the MPEG mapping into the Generic Container. The value of byte 14 shall be set to 0Fh for the AVC NAL unit stream and 10h for the AVC byte stream.

Notes:

1 The custom wrappings (byte 16 values of 03h, 06h, 07h and 08h) are as defined in SMPTE ST 381-2 which inherits the definition from SMPTE ST 381-1.

2 Essence Container Label version is as follows:

For byte 16 values 01, 02, 03, 06, 07 and 08h, the correct version is 0Ah.
For byte 16 value 09h, the correct version is 0Dh.

3 A previous published version of this document included the value 04h for byte 16, meaning Custom: PES Wrapping. The labels to allow this wrapping were never registered and it is no longer a supported value.

8.2 Picture Essence Coding Label

The Picture Essence Coding Label is used in the Generic Picture Essence Descriptor. The values for the Picture Essence Coding Label for the AVC are given in Table 3. These Label values are listed in the SMPTE Labels Register.

Table 3 – Specification of the Picture Essence Coding Label

Byte No.	Description	Value (hex)	Meaning
1-8	Registry Designator	See SMPTE ST 400	Designator value as defined in SMPTE ST 400
9	Parametric	04h	Node used to define parametric data
10	Picture Essence	01h	Identifies picture essence coding
11	Picture Coding Characteristics	02h	Identifies picture coding characteristics
12	Compressed Picture Coding	02h	Identifies compressed picture coding
13	MPEG Picture Coding	01h	Identifies MPEG picture coding
14	AVC Picture Coding and Profile Category	3xh	Identifies AVC picture coding by setting bits [7..4] to the value of '3h' as listed in SMPTE Labels Register Identifies Profile category for AVC using bits [3..0] as listed in SMPTE Labels Register
15	AVC Profile Constraints	xxh	Identifies Profile and its constraints for AVC as listed in SMPTE Labels Register
16	AVC Coding variants	yyh	Identifies variants of AVC coding as listed in SMPTE Labels Register

Byte 14 identifies the AVC picture coding and Profile category and shall be set to 31h for Predictive profiles and to 32h for Intra profiles.

Byte 15 identifies the Profile and its constraints of the AVC picture coding. Bits [7..4] shall indicate the Profile. A value of '0h' of bits [3..0] shall indicate a generic application that has no coding constraints. Other values of bit [3..0] shall be defined by applications defining specific coding constraints and registered in SMPTE Labels Register.

Byte 16 identifies details of the AVC coding variants. The values of byte 16 shall be defined depending on the value of byte 15, and shall be registered in SMPTE Labels Register as necessary to identify particular applications in other SMPTE Engineering Documents that define specific coding constraints. When bits [3..0] of byte 15 indicates a generic application, the value of byte 16 shall be set to '01h'.

The values for the Picture Essence Coding Label and the labels for a generic application are given in Table 4.

Table 4 – AVC Picture Essence Coding Labels

Node/ Leaf	UL	Meaning
Node	06.0E.2B.34.04.01.01.0A 04.01.02.02.01.30.00.00	H.264/MPEG-4 AVC Video
Node	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.00.00	H.264/MPEG-4 AVC Predictive Profiles
Node	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.10.00	H.264/MPEG-4 AVC Baseline Profiles
Leaf	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.10.01	H.264/MPEG-4 AVC Baseline Profile Unconstrained Coding
Node	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.11.00	H.264/MPEG-4 AVC Constrained Baseline Profiles
Leaf	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.11.01	H.264/MPEG-4 AVC Constrained Baseline Profile Unconstrained Coding
Node	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.20.00	H.264/MPEG-4 AVC Main Profile
Leaf	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.20.01	H.264/MPEG-4 AVC Main Profile Unconstrained Coding
Node	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.30.00	H.264/MPEG-4 AVC Extended Profile
Leaf	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.30.01	H.264/MPEG-4 AVC Extended Profile Unconstrained Coding
Node	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.40.00	H.264/MPEG-4 AVC High Profile
Leaf	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.40.01	H.264/MPEG-4 AVC High Profile Unconstrained Coding
Node	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.50.00	H.264/MPEG-4 AVC High 10 Profile
Leaf	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.50.01	H.264/MPEG-4 AVC High 10 Profile Unconstrained Coding
Node	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.60.00	H.264/MPEG-4 AVC High 422 Profile
Leaf	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.60.01	H.264/MPEG-4 AVC High 422 Profile Unconstrained Coding
Node	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.70.00	H.264/MPEG-4 AVC High 444 Predictive Profile
Leaf	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.31.70.01	H.264/MPEG-4 AVC High 444 Predictive Profile Unconstrained Coding
Node	06.0E.2B.34.04.01.01.0A 04.01.02.02.01.32.00.00	H.264/MPEG-4 AVC Intra Profiles
Node	06.0E.2B.34.04.01.01.0A 04.01.02.02.01.32.20.00	H.264/MPEG-4 AVC High 10 Intra Profile
Leaf	06.0E.2B.34.04.01.01.0A 04.01.02.02.01.32.20.01	H.264/MPEG-4 AVC High 10 Intra Profile Unconstrained Coding
Node	06.0E.2B.34.04.01.01.0A 04.01.02.02.01.32.30.00	H.264/MPEG-4 AVC High 422 Intra Profile

Leaf	06.0E.2B.34.04.01.01.0A 04.01.02.02.01.32.30.01	H.264/MPEG-4 AVC High 422 Intra Profile Unconstrained Coding
Node	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.32.40.00	H.264/MPEG-4 AVC High 444 Intra Profile
Leaf	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.32.40.01	H.264/MPEG-4 AVC High 444 Intra Profile Unconstrained Coding
Node	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.32.50.00	H.264/MPEG-4 AVC CAVLC 444 Intra Profile
Leaf	06.0E.2B.34.04.01.01.0D 04.01.02.02.01.32.50.01	H.264/MPEG-4 AVC CAVLC 444 Intra Profile Unconstrained Coding

Notes:

1 The node entries do not constitute valid essence coding UL values that can be used in the descriptor.

2 In Table 4 and the corresponding entries in SMPTE Labels Register, the term "Unconstrained Coding" is taken to mean "Profile compliant coding with no additional constraints."

9 Essence Descriptor for AVC Mapping

9.1 AVC Sub Descriptor

For the AVC wrapping, the AVC Sub Descriptor, which is strongly referenced from the CDCI Descriptor or the RGBA Descriptor, should be used. The AVC Sub Descriptor consists of the AVC-specific properties as shown in Table 5. If the properties in Table 5 are present, their values shall apply to the entire Essence described. Therefore, these properties allow fixed picture patterns to be identified.

Note: In some legacy cases, the MPEG Video Descriptor has been used to describe AVC instead of the AVC Sub Descriptor.

Table 5 – AVC Sub Descriptor

Item Name	Type	Len	Item UL	Req ?	Meaning	Default
AVC Sub Descriptor	Set UL	16	See Table 9	Req	Defines the AVC Sub Descriptor Set	
Length	BER Length	4		Req	Set length	
All items from the Sub Descriptor set (except for the Group UL) as specified in SMPTE ST 377-1						
AVC Decoding Delay	UInt8	1	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.0E.00.00	Req	Specifies the delay required for decoded pictures in number of access units. The value is given by the presentation time of the first presented picture in a GOP minus the decoding time of the first decoded picture in the GOP. The value shall be set to zero if there are no B Pictures in the essence stream. The value shall be set to FFh if the delay is unknown.	
AVC Constant B Picture Flag	Boolean	1	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.03.00.00	Opt	TRUE if the number of the consecutive B Pictures is always constant. FALSE if the number of the consecutive B Pictures is not constant or is unknown.	FALSE

Item Name	Type	Len	Item UL	Req ?	Meaning	Default
AVC Coded Content Kind	UInt8Enum	1	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.04.00.00	Opt	'0' = "Unknown" '1' = "Progressive Frame Picture" '2' = "Interlaced Field Picture" '3' = "Interlaced Frame Picture" '4' = "Interlaced Frame and Field Picture" Enumerated value specifying Picture type and Coding type. See Table 6 for details,	0
AVC Closed GOP Indicator	Boolean	1	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.06.00.00	Opt	TRUE if all GOPs are started with an IDR Picture. FALSE if not all GOPs are started with an IDR Picture or if it is unknown whether all GOPs start with an IDR Picture	FALSE
AVC Identical GOP Indicator	Boolean	1	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.07.00.00	Opt	TRUE if every GOP in the sequence has the same number of pictures and the same types of pictures in the same order. FALSE if not all GOPs are known, or if there is at least one GOP in the sequence which has any of the following: 1. a different number of pictures 2. the same type of pictures but in a different order 3. different types of pictures.	FALSE
AVC Maximum GOP Size	UInt16	2	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.08.00.00	Opt	Specifies the maximum occurring spacing between I Pictures. A value of 0 or the absence of this property indicates either there is no limit to the maximum GOP or the maximum GOP is unknown. If an encoder does not know the value, this property shall not be present.	
AVC Maximum B Picture Count	UInt16	2	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.09.00.00	Opt	Specifies the maximum number of B Pictures between P or I Pictures. If an encoder does not know the value, this property shall not be present.	
AVC Maximum Bitrate	UInt32	4	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.0B.00.00	Opt	Maximum bit rate of the AVC stream in bit/s is given by bit_rate_scale and bit_rate_value_minus1 in the HRD parameters in the sequence parameter set. The equivalent value is assigned for this property of the stream even if the stream does not include the HRD parameters. If an encoder does not know the value, this property shall not be present.	
AVC Average Bitrate	UInt32	4	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.14.00.00	Opt	Average bitrate of the AVC stream in bit/s over the entire AVC bitstream. If an encoder does not know the value, this property shall not be present.	
AVC Profile	UInt8	1	06.0E.2B.34. 01.01.01.0E. 04.01.06.06.	Opt	Specifies the AVC video profile. The value is taken from profile_idc in the sequence parameter set.	

Item Name	Type	Len	Item UL	Req ?	Meaning	Default
			01.0A.00.00			
AVC Profile Constraint	UInt8	1	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.0C.00.00	Opt	Specifies the AVC video profile constraint flags. The value of bit [7..2] is taken from constraint_set0_flag, constraint_set1_flag, constraint_set2_flag, constraint_set3_flag. constraint_set4_flag and constraint_set5_flag in the sequence parameter set. The value of bit [1..0] shall be set to zero.	
AVC Level	UInt8	1	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.0D.00.00	Opt	Specifies the AVC level. The value is taken from level_idc in the sequence parameter set.	
AVC Maximum Ref Frames	UInt8	1	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.0F.00.00	Opt	Specifies the maximum number of reference frames. The value is the maximum value of max_num_ref_frames within all sequence parameter sets.	
AVC Sequence Parameter Set Flag	UInt8	1	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.10.00.00	Opt	Specifies the location and the constancy of sequence parameter sets See Table 7 for details.	0
AVC Picture Parameter Set Flag	UInt8	1	06.0E.2B.34. 01.01.01.0E. 04.01.06.06. 01.11.00.00	Opt	Specifies the location and the constancy of picture parameter sets See Table 8 for details	0

Table 6 – Meaning of AVC Coded Content Kind

Value	Source image	Picture type	Coding type
0	Unknown		
1	Progressive	Frame	Frame Coding
2	Interlaced	Field	Field Coding
3	Interlaced	Frame	Frame Coding including Macroblock Adaptive Frame-Field Coding
4	Interlaced	Frame and Field	Picture Adaptive Frame-Field Coding Frame Coding may use Macroblock Adaptive Frame-Field Coding
5 – Fh			Reserved

Notes:

1 In the case of Value 1, 3 and 4 streams, if byte 15 of the Picture Element Key is 05h (Frame Wrapped), then byte 16 of the AVC Essence Container Label is 01h (Frame Wrapping).

2 For a Value 2 stream, if byte 15 of the Picture Element Key is 05h (Frame Wrapped), then byte 16 of the AVC Essence Container Label is either 01h (Frame Wrapping) or 09h (Field Wrapping) as defines in Section 7.3.

Table 7 – Meaning of AVC Sequence Parameter Set Flag

Bit number	Name	Values and Descriptions	Default
7	Constancy flag	Specifies whether all sequence parameter sets are constant 0: unknown (sequence parameter sets may vary in the stream) 1: constant (all sequence parameter sets are identical throughout the stream)	0
6-4	In-band location	Specifies the location of sequence parameter sets in the stream 0: Unknown or no specific location 1: Only the first access unit in the stream 2: Every access unit in the stream 3: Periodically placed at the first access unit in each GOP others: reserved	0
3-0	reserved	Reserved for future use and shall be set to zero	0

Table 8 – Meaning of AVC Picture Parameter Set Flag

Bit number	Name	Values and Descriptions	Default
7	Constancy flag	Specifies whether all picture parameter sets are constant 0: unknown (picture parameter sets may vary in the stream) 1: constant (all picture parameter sets are identical throughout the stream)	0
6-4	In-band location	Specifies the location of picture parameter sets in the stream 0: Unknown or no specific location 1: Only the first access unit in the stream 2: Every access unit in the stream 3: Periodically placed at the first access unit in each GOP others: reserved	0
3-0	reserved	Reserved for future use and shall be set to zero	0

9.2 Key for the AVC Sub Descriptor

The key (UL) for this Local Set shall be as defined in Table 9.

Table 9 – Key for AVC Sub Descriptor

Byte No.	Description	Value (hex)	Meaning
1-13	Defined in the Structural Header Metadata Implementation section of SMPTE ST 377-1		
14	Set Kind (1)	01_h	AVC Sub Descriptor
15	Set Kind (2)	$6E_h$	
16	Reserved	00_h	Reserved

Note: The AVC Sub Descriptor Key version is 01h.

10 Index Table for AVC Mapping

This section clarifies how the Index Entry Properties are set for indexing an AVC stream.

There are several properties in the Index Table Entry that have specific meanings for an AVC Long GOP MPEG Index Table. The Flags Property, shown in Table 10, shall be correctly set for AVC according to the text below Table 10 that describes the conditions for setting the Flags.

Note: Flag bits 3-0 are reserved for use in SMPTE Essence mapping specifications in SMPTE ST 377-1 and the definitions for AVC Essence mapping are given in this document.

Table 10 – Index Table Entry Properties for AVC

Parameter	Type	Meaning	Use
Temporal Offset	Int8	Offset in Edit Units from Display Order to Coded Order (see SMPTE ST 377-1 for usage)	Used to find the Index Entry for a stored Picture given its Display Temporal position.
Key Frame Offset	Int8	Offset in Edit Units to previous Key Picture. The value is zero if this is a Key Picture.	The offset to the Key Picture
Flags	EditUnitFlag	Flags for this Edit Unit Bit 7: Random Access Bit 6: Sequence Parameter Set Bits 5,4: Forward/backwards prediction flag Bit 3: Offsets out of range Bit 2: Supplementary Picture type flag Bits 1,0: AVC picture type	Used to fetch a sufficient number of frames to ensure that all Key Pictures required to decode any Picture within the MPEG GOP sequence are available

Temporal Offset: This is an offset used to allow lookups in the Index Table. It corresponds to the difference between the display order and coded order for the indexed picture measured in Edit Units, not in frames. This is explained in the text of SMPTE ST 377-1.

If the numerical range of this Property is exceeded, then bit 3 of the Flags shall be set, and this Property shall be clipped to the maximum value which can be represented. The actual temporal offset may be determined by inspecting Index Entries for neighboring pictures until a valid temporal offset is found. The value of temporal offset may be determined by the number of Index Entries inspected, their picture types and the valid temporal offset value found.

Key Frame Offset: This is the offset measured in Edit Units (not frames) to the Key Picture (I Picture) required for decoding the indexed picture.

If the stream consists of I Pictures only, then this parameter shall be set to zero since there would be no Key Pictures required to decode an indexed picture.

If the numerical range of this Property is exceeded, then bit 3 of the Flags shall be set, and this Property shall be clipped to the maximum value which can be represented. The actual Key Frame Offset may be determined by inspecting Index Entries for neighboring pictures until a valid Key Frame Offset is found. The value of Key Frame Offset may be determined by the number of Index Entries inspected, their picture types and the valid Key Frame Offset value found.

Key Frame Offset is always negative for AVC streams. The offset is measured in Edit Units and is the coded order offset of the Key Picture Index Entry. This keeps the meaning of the sign of Key Frame Offset consistent with the sign of temporal offset. It also means that the only display order process is to use the temporal offset to find the appropriate index entry. Index entries are always stored in coded order.

Flag Bit 7: This is the random access bit when a random access point in the AVC Long GOP stream is encountered. A random access point is one where decoding can commence at that point in the stream with valid parameter sets.

All random access points will be the first access unit in a GOP.

All Pictures needed for display after the recovery point will have no decoding dependency on any data preceding the random access point. They are also present in the decoding sequence after this random access point.

Flag Bit 6: This bit shall mark a sequence parameter set in the AVC stream.

Flag Bits 5 and 4: These bits shall be set to indicate the temporal dependence of the Indexed Picture. These bits may be set naïvely depending on Picture type or may be set strictly by depending on prediction modes. The naïve settings shall be used by an encoder when strict setting is not possible. The naïve and strict settings of these bits by an encoder shall be as follows:

Naïve setting of bits 5 and 4:

- 00 I Picture
- 10 P Picture
- 11 B Picture

Strict setting of bits 5 and 4:

- 00 No prediction. This is always the case for I Pictures.
- 10 Forward prediction from previous Picture.
- 01 Backward prediction to future Picture.
- 11 Forward and backward prediction.

Flag Bit 3: This bit shall indicate that the numerical range of the Temporal Offset or Key Frame Offset fields has been exceeded. In this case, the offset and/or Key Picture may be determined by inspecting other Index Entries to determine the precise values.

Flag Bit 2: This bit indicates supplementary information on each Picture type. This bit shall be set to 1 for IDR Picture for I Picture and referenced B Picture (Br Picture) for B Picture. It is recommended that this bit be set to 1 for reference for P Picture.

Flag Bits 1 and 0: These bits indicate AVC picture type. The setting of these bits shall be as follows:

- 00 I Picture
- 10 P Picture
- 11 B Picture

Table 11 – Picture type identification by Flag Bits [2..0]

Flag Bits [2..0]	Picture type
000	non-IDR I Picture
100	IDR Picture
010	P Picture This Picture may be referenced.
110	referenced P Picture
011	non-referenced B Picture
111	referenced B Picture(Br Picture)
others	Not used

Annex A Bibliography (Informative)

Note: All references in this document to other SMPTE documents use the current numbering style (e.g. SMPTE ST 336:2007) although, during a transitional phase, the document as published (printed or PDF) may bear an older designation (such as SMPTE 336M-2007). Documents with the same root number (e.g. 336) and publication year (e.g. 2007) are functionally identical.

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

SMPTE ST 381-1:2005, Television — Material Exchange Format (MXF) — Mapping MPEG Streams into the MXF Generic Container

SMPTE ST 400:2012, SMPTE Labels Structure

SMPTE Labels Register <https://smpte-ra.org/smpte-metadata-registry>

Annex B Property Values of Picture Essence Descriptors (Informative)

This Annex provides clarification of the “Req” and “B.Effort” property values of the Picture Essence Descriptors for compatibility. The property values are implemented as shown in Table 12.

Table 12 – “Req” and “B.Effort” Property values for AVC

Item Name	Type	Len	Meaning, Appropriate values
File Descriptor			
Sample Rate	Rational	8	Equal to Edit Rate Gives Frame Rate where Frame or Field encoded stream is frame-wrapped. Gives Field Rate where Field encoded stream is field-wrapped.
Essence Container	UL	16	See Table 2
Generic Picture Essence Descriptor			
Frame Layout	UInt8	1	Scanning of the source picture I.e. 0 (FULL_FRAME): Progressive pictures 1 (SEPARATE_FIELDS): Interlaced pictures
Stored Width	UInt32	4	Horizontal size of the Stored Rectangle E.g. 1280: 1280x720 image structure 1440: 1440x1080 image structure 1920: 1920x1080 image structure
Stored Height	UInt32	4	Vertical size of the Stored Rectangle In cases where additional lines compared to the Sampled Rectangle are being used, e.g. padding for macro-block alignment, these additional lines are included in the Stored Rectangle. E.g. 544: 1080-line interlaced 720: 720-line progressive 1088: 1080-line progressive
Aspect Ratio	Rational	8	E.g. 4:3, 16:9
Video Line Map	Array of Int32	8+8	E.g. (21, 584): 1080-line interlaced (26): 720-line progressive (42): 1080-line progressive
CDCI Picture Essence Descriptor			
Component Depth	UInt32	4	The number of active bits per component sample. E.g. 10: 10bit
Horizontal Subsampling	UInt32	4	The horizontal subsampling factor of the color difference samples relative to the luma samples in the Stored Rectangle. E.g. 1: 4:4:4 2: 4:2:2, 4:2:0

Item Name	Type	Len	Meaning, Appropriate values
RGBA Picture Essence Descriptor			
Pixel Layout	RGBA Layout	16	A property of type RGBALayout, whose value specifies the type, order and size of the components within the Pixel as defined in SMPTE ST 377-1 Annex G.
AVC Sub Descriptor			
AVC Decoding Delay	UInt8	1	See Table 5

Annex C Possible Structure of AVC Long GOP (Informative)

This Annex describes a possible structure of AVC Long GOP streams for MXF wrapping to provide fast access to target Pictures and reduce decoder complexity.

- All P Pictures consist of only P-Slices and all B Pictures or Br Pictures consist of only B-Slices.
- The first access unit in coded order is an IDR or non-IDR I Picture in GOP and a sequence parameter set and a picture parameter set are present in the first access unit of each GOP.
- For each picture in the GOP, the corresponding Key Picture is present within the same GOP or the immediately preceding GOP.
- A P Picture does not refer to a B Picture.
- The coded order for all I- and/or P Pictures is the same as their display order.
- The coded order for non-reference B Pictures that are contiguous in the bit stream is the same as their display order.

Figure 1 shows an example of the constrained structure of an AVC Long GOP stream.

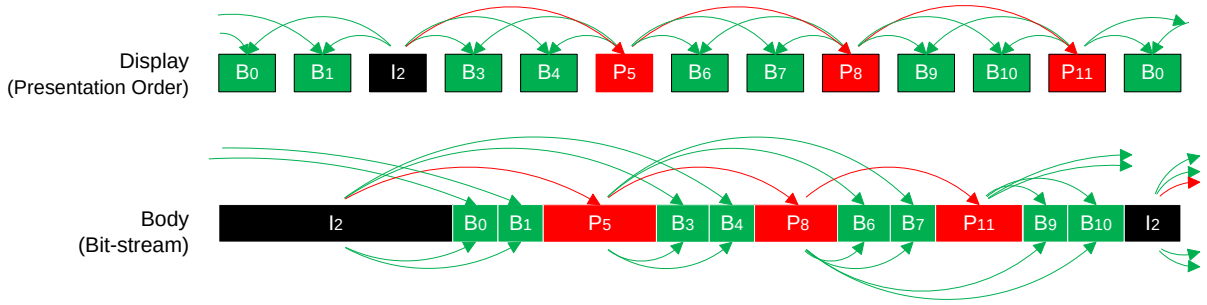

Figure 1 – Example of Constrained AVC Long GOP structure

Figure 2 shows another example of the constrained structure of a “Hierarchical coding structure” AVC Long GOP stream. Note that in AVC B Pictures can be reference pictures, and these behave in a similar way to P Pictures in MPEG-2.

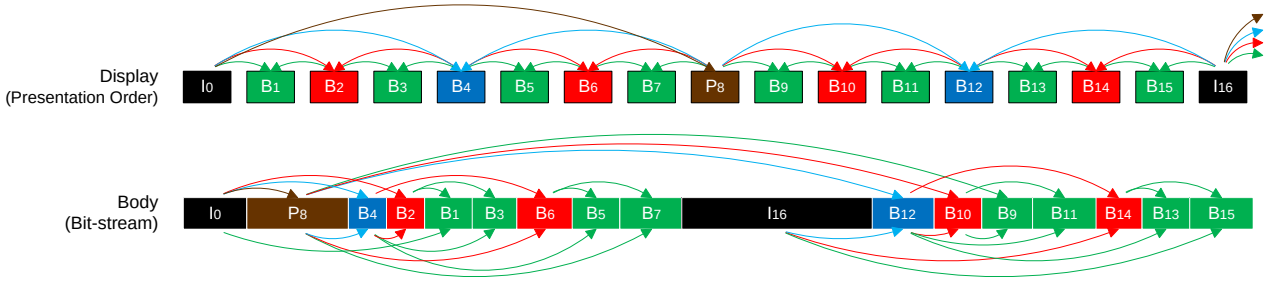

Figure 2 – Example of Hierarchical coding structure