SMPTE STABLE DOCUMENT

The attached SMPTE Engineering Document has been declared "Stable" by the controlling Technology Committee.

The SMPTE Operations Manual for Standards states:

> *A document should be stabilized if it is believed to be substantially correct, does not contain harmful or misleading recommendations, may still be relevant to equipment or practices in use, is stable, but does not represent current technology, and need not be subject to future reviews.*
>
> *A Stable document shall still be made available and offered for sale by the Society, but it shall be prefaced by a cover page explaining its current status.*
>
> *At any time, a Technology Committee may revise, amend, or otherwise initiate a new Project on a Stable document.*

A Stable document is "In Force", and not deprecated or withdrawn.

* * * * * * * * * * *

Note:

SMPTE "Stable" documents were previously described as "Archived" and the attached document may be marked as "Archived". The status of a SMPTE document described as "Archived" is exactly as described above for a "Stable" document.

Stable documents may not adhere to the latest style and format of SMPTE documents, or to current usage of normative language. Suitable care should be taken in interpretation.

SMPTE STANDARD

Mapping VC-1 into the MXF Generic Container

Table of Contents

Approved
May 4, 2009

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 Standard 2037 was prepared by Technology Committee 31FS.

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 defines how VC-1 bitstreams shall be mapped in the MXF Generic Container (SMPTE 379M).

VC-1 defines three possible profiles for compliant bitstreams: Simple, Main and Advanced. All of these profiles are supported by this standard.

VC-1 bitstreams comprise VC-1 encoded Pictures and both intra-coding and inter-coding mechanisms can be used. Both coding mechanisms are supported in this standard.

This standard maps the VC-1 encoded Pictures as either frame-wrapped or clip-wrapped using the mechanisms defined in SMPTE 379M, and both progressive and interlaced Picture representations are supported in this standard.

This standard defines the KLV coding, the Essence Container Universal Label, the Essence Compression Universal Label and the Essence Descriptor to be used when wrapping VC-1 bitstreams in the MXF Generic Container.

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.

SMPTE 377-1-2009, Material Exchange Format (MXF) — File Format Specification

SMPTE 379M-2004, Television — Material Exchange Format (MXF) — MXF Generic Container

SMPTE 400M-2004, Television — SMPTE Labels Structure

SMPTE 421M-2006, Television — VC-1 Compressed Video Bitstream Format and Decoding Process

Amendment 1-2007 to SMPTE 421M-2006

4 Glossary of Acronyms, Terms and Data Types

The general glossary of acronyms, terms and data types used in the MXF specification is given in SMPTE 377-1 and is supplemented in SMPTE 379M. Definitions of terms, abbreviations and symbols relating to VC-1 are given in SMPTE 421M.These glossaries are not repeated here to avoid any divergence of meaning.

5 Overview

VC-1 bitstreams may make use of both intra-coding and inter-coding mechanisms. For this reason, each picture can be coded as an I-picture, P-picture, skipped picture, BI-picture, or as a B-picture. Additionally, each Picture can be represented as a progressive picture or as a pair of encoded fields.

Throughout this document, the sequence of bytes in the VC-1 bitstream used to represent an encoded Picture, shall be called Data Access Unit, regardless of the specific encoding scheme.

6 MXF Encoders — Locating Data Access Units

This section defines how MXF encoders shall locate Data Access Units in the VC-1 bitstreams when wrapping VC-1 encoded Pictures in the MXF Essence Container.

6.1 Advanced Profile

In the case of the Advanced Profile, a Data Access Unit comprises all the coded data for a picture and any flushing bits that follows it, up to but not including the start code of the next access unit (or any stuffing bytes before it). The end-of-sequence start code, if present, shall be included in the Data Access Unit of the picture immediately preceding this start code. The start of the next Data Access Unit shall be the byte of either a sequence start code, an entry point start code or a frame start code.

Also, if the frame is not preceded by a sequence start code and a sequence header or an entry point start code and an entry point header, the Data Access Unit shall begin with a frame start code. Otherwise, the Data Access Unit shall start with the first byte of the first of these structures (excluding any stuffing bytes) before the frame start code. A Data Access Unit shall also include any user data start code and user data bytes at the sequence, entry point, frame or field level.

6.2 Simple Profile and Main Profile

In the case of Simple Profile or Main Profile, and in accordance with Section 7.1 in SMPTE 421M, the location of each coded Picture and its size shall be communicated by the transport layer.

For this reason, the way MXF encoders obtain such information when wrapping VC-1 bitstreams in MXF is out of the scope of this standard. Nevertheless, once decoders use that information to wrap the Pictures in MXF, decoders can easily locate Pictures in the EssenceContainer as defined in the following section.

7 Data Access Units in the MXF Generic Container

This mapping shall use the MXF Generic Container in either the frame-based wrapping or clip-based wrapping mode defined in SMPTE 379M. Knowledge of the MXF Generic Container terminology is assumed throughout this document.

7.1 Frame Wrapping

For Frame wrapping, the sequence of Data Access Units of a VC-1 bitstream shall be placed in a sequence of Picture Elements, placed in a sequence of Content Packages. Each Content Package shall have only one Picture Element of a given bitstream.

The VC-1 Data Access Units may optionally be interleaved with other essence components in the frame-wrapped essence container. These other essence components shall be as defined by this or other MXF mapping standards. All essence element types shall be frame wrapped.

In this kind of wrapping, MXF decoder can easily identify each Data Access Unit, since each Data Access Unit shall be independently KLV wrapped as a Picture Element.

7.2 Clip Wrapping

For Clip wrapping, the sequence of Data Access Units of a VC-1 bitstream shall be placed in a single Picture Element, in a Content Package of a MXF Generic Container Content Package.

The clip-wrapped VC-1 Picture Element may be the sole component in the MXF Generic Container Content Package. The clip-wrapped VC-1 Picture Element may also be used in the MXF Generic Container Content Package in sequence with other clip-wrapped Picture Elements.

In this kind of wrapping Data Access Units are not independently KLV wrapped. Therefore, if Index Tables indexing the full extent of the Essence Container are present, these may be used to locate Data Access Units. Otherwise:

- for Advanced Profile VC-1 bitstreams, the Data Access Unit location method in Section 6.1 above shall be used;

- for Simple and Main profiles, alternative mechanisms need to be defined enable decoders to locate individual Data Access Units in the MXF Generic Container; the Picture Essence Coding property of the Generic Picture Essence Descriptor shall identify the mechanism to be used.

MXF Encoders are strongly encouraged to make use of Index Tables in order to reduce decoder complexity.

8 Key-Length-Value Coding

8.1 Essence Element Key

The values of the first 12 bytes of the essence element Key are defined in SMPTE 379M. The version (byte 8) shall be set to 0x01. The values of the last four bytes of the Picture Element Key are given in Table 1.

Table 1 – Key Value for the VC-1 Picture Element

Byte No.	Description	Value (hex)	Meaning
1~12	Defined in SMPTE 379M		See SMPTE 379M
13	Item Type Identifier	15h	GC Picture Item (as defined in SMPTE 379M)
14	Essence Element Count	kkh	Count of Picture Elements in the Picture item
15	Essence Element Type	0Ah 0Bh	Frame-wrapped VC-1 Picture Element Clip-wrapped VC-1 Picture Element
16	Essence Element Number	Nnh	The Number (used as an Index) of this Picture Element in the Picture Item

8.1.1 Essence Element Count – Byte 14

This is a count of the number of Picture Elements in the Picture Item of the Generic Container.

8.1.2 Essence Element Type – Byte 15

The value of 0Ah identifies that each VC-1 codestream is frame-wrapped. The value of 0Bh identifies that the sequence of VC-1 codestreams are clip-wrapped.

8.1.3 Essence Element Number – Byte 16

This is a number used as an index to identify this instance of the Element Type within the Picture Item. Each Element within an Item shall have a unique value between 00h and 7Fh, as defined by SMPTE 379M, which shall remain constant within the Generic Container.

8.2 Length

The length field shall comply with SMPTE 379M Section 5.5.1. The length field should be 4 byte BER long-form encoded (i.e. 83h.xx.yy.zz) for frame-based wrapping and should be 8 byte BER long-form encoded (i.e. 87h.aa.bb.cc.dd.ee.ff.gg) for clip-based wrapping.

Note: Although the use of 8-byte BER long-form encoding for clip-based wrapping will be sufficient for all foreseen applications, the use of 4-byte BER long-form encoded length fields limits the size to 16Mbytes. This may be insufficient for very high resolution pictures in which case longer length fields should be used.

8.3 Value

8.3.1 Frame-wrapped

The value field shall comprise a single VC-1 Data Access Unit.

8.3.2 Clip-wrapped

The value field shall comprise a sequence of one or more concatenated VC-1 Data Access Units.

9 SMPTE Universal Labels for VC-1

9.1 Essence Container UL

The values for the Essence Container UL are given in Table 2.

Table 2 – Specification of the Essence Container Label

Byte No.	Description	Value (hex)	Meaning
1-12	Defined by Generic Container		As defined in SMPTE 379M
13	Essence Container Kind	02h	MXF Generic Container
14	Mapping Kind	12h	VC-1 Picture Element (as listed in the register defined by SMPTE 400M (RP 224)
15	Content Kind	01h 02h	Frame-wrapped Picture Element Clip-wrapped Picture Element
16	Reserved	00h	

The Essence Container UL is used within a batch of ULs in Partition Packs and the Preface set and on its own in the Essence Descriptor. These UL values are listed in the register defined by SMPTE 400M (RP 224).

9.2 Picture Essence Coding UL

Values for the Picture Essence coding UL are given in Table 3.

Table 3 – Specification of the Picture Essence Compression Label

Byte No.	Description	Value (hex)	Meaning
1-8	Registry Designator	See SMPTE 400M	Designator value is defined in SMPTE 400M
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 specialized compression
13		04h	Identifies VC-1 picture coding
14		xxh	Identifies VC-1 Profile and Level according to the table below
15		00h	Unused
16		00h	Unused

The Picture Essence Coding UL is used in the Generic Picture Essence Descriptor. This UL is listed in the register defined by SMPTE 400M (RP 224).

The correspondence between byte 13 and the VC-1 Profile and Level described in SMPTE 421M is shown in Table 4.

Table 4 – Correspondence Between Byte 13 and VC-1 Profile Level

Profile	Level	Abbreviation	Value of Byte 13
Simple	Low	SP@LL	01h
	Medium	SP@ML	02h
Main	Low	MP@LL	03h
	Medium	MP@ML	04h
	High	MP@HL	05h
Advanced	L0	AP@L0	06h
	L1	AP@L1	07h
	L2	AP@L2	08h
	L3	AP@L3	09h
	L4	AP@L4	0Ah

10 VC-1 Picture Essence Descriptor

This section defines the VC1PictureEssenceDescriptor that extends the CDCI Picture Essence Descriptor defined in SMPTE 377-1 with additional properties specific to VC-1.

All VC-1 bitstreams wrapped in the MXF Essence Container according to this standard shall be described in the MXF Header Metadata with a VC1PictureEssenceDescriptor according to the rules defined in SMPTE 377-1 on how to link such descriptors with the Essence Container, and as shown in Table 5.

Table 5 – VC-1 Picture Essence Descriptor

Item Name	Type	Len	UL Value	Req ?	Meaning	Default
VC-1 Picture Essence Descriptor	Set UL	16	See table below	Req	Defines the VC-1 Picture Essence Descriptor Set	
Length	BER Length	4		Req	Set length	
All items from the MXF Format CDCI Descriptor						
VC1InitializationMetadata	Datastream	16	06 0e 2b 34 01 01 01 0C 04 01 06 04 01 00 00 00	D/Req	See Section 10.1.	
VC1SingleSequence	Boolean	1	06 0e 2b 34 01 01 01 0C 04 01 06 04 02 00 00 00	Opt	TRUE if the essence consists of a single VC-1 sequence. False if there are a number of sequences. This flag implies that the sequence header information is not varying in the essence stream.	Unknown
VC1CodedContentType	Enum	1	06 0e 2b 34 01 01 01 0C 04 01 06 04 03 00 00 00	Opt	0= "Unknown" 1= "Progressive" 2= "Interlaced" 3= "Mixed" an enumerated value which tells if the underlying content which was VC-1 coded was of a known type	Unknown

Item Name	Type	Len	UL Value	Req ?	Meaning	Default
VC1IdenticalGOP	Boolean	1	06 0e 2b 34 01 01 01 0C 04 01 06 04 04 00 00 00	Opt	TRUE if every GOP in the sequence is constructed the same.	FALSE
VC1MaxGOP	UInt16	2	06 0e 2b 34 01 01 01 0C 04 01 06 04 05 00 00 00	Opt	Specifies the maximum occurring spacing between I frames A value of 0 or the absence of this property implies no limit to the maximum GOP	no limit
VC1BPictureCount	UInt16	2	06 0e 2b 34 01 01 01 0C 04 01 06 04 06 00 00 00	Opt	Specifies the maximum number of B pictures between P or I frames. BI-pictures count as B-pictures and skipped-pictures count as P-pictures	
VC1AverageBitRate	UInt32	4	06 0e 2b 34 01 01 01 0C 04 01 06 04 07 00 00 00	Opt	Average bit rate of VC-1 video elementary stream in bit/s.	
VC1MaximumBitRate	UInt32	4	06 0e 2b 34 01 01 01 0C 04 01 06 04 08 00 00 00	Opt	Maximum bit rate of VC-1 video elementary stream in bit/s.	
VC1Profile	UInt8	1	06 0e 2b 34 01 01 01 0C 04 01 06 04 09 00 00 00	Opt	Specifies the VC-1 video profile Coded as per SMPTE 421M Annex J.1.1	
VC1Level	UInt8	1	06 0e 2b 34 01 01 01 0C 04 01 06 04 0a 00 00 00	Opt	Specifies the VC-1 video level. Coded as per SMPTE 421M Annex J.1.2	

10.1 VC1InitializationMetadata

This property encodes the Metadata defined in SMPTE 421M Annex J and shall be present for Simple and Main Profile bitstreams. The value of the property shall consist of the serialization of the STRUCT_SEQUENCE_HEADER_B as defined in Table 261 of SMPTE 421M (96 bits), followed by the STRUCT_SEQUENCE_HEADER_C as defined in Table 263 of SMPTE 421M (32 bits).

10.2 Key Value

The set Key of the VC-1 Picture Descriptor shall be as defined in Table 6.

Table 6 – Key Value for the VC-1 Picture Essence Descriptor

Byte No.	Description	Value	Meaning
1~13	As defined in SMPTE 377-1, Common Key Value for Structural Metadata (Table 13)		Values for all MXF structural metadata sets
14~15	Set Kind	01.5Fh	Defines the Key value for the VC-1 Picture Essence Descriptor
16	Reserved	00h	Reserved value

10.3 Length Value

The set length shall be BER long form encoded. It is preferred that the length field uses 4 bytes.

11 Pre-Change

In some applications, the first relevant Picture stored in a MXF file with VC-1 data is a inter-coded picture. An example is partial restore operation where a MXF file is created by selecting a portion of original VC-1 encoded material, and where the first relevant picture happens to be a B-picture. However, in order to decode that inter-coded Picture, decoders need to be fed with Pictures that precede in the bitstream the required relevant Picture. This process is commonly referred to as decoder pre-charge.

This section presents the mechanism to be used in order to allow users to specify a inter-coded picture as the first relevant picture, while still wrapping as well all Pictures needed for decoder pre-charge.

The MXF Generic Container shall include Data Access Units needed to guarantee that the stored Picture data starts and ends at GOP boundaries. The portion of that bitstream to actually be considered relevant by applications shall be specified in the MXF Header Metadata as follows:

- The Essence Container is represented by a Top-Level Source Package;
- The track within the Essence Container where the VC-1 bitstream is wrapped is represented by a TimelineTrack in the SourcePackage
- The Origin property of that TimelineTrack shall indicate the number of Edit Units, which equals the number of Data Access Units, to skip from the start of that Track onwards until the first Picture which is relevant for the content is found.

Some of the pictures at the end of the MXF Generic Container may not be relevant to the content either for the same reasons presented above. However, these Pictures may also be necessary for the decoding of relevant pictures. Following the reasoning above, the Duration parameter in the TimelineTrack shall be set accordingly in order to clearly signal the last relevant picture, which may not be the last picture KLV encoded in the MXF Generic Container.

12 Index Tables

SMPTE 377-1 describes how MXF Index Tables can be used to perform temporal re-ordering for MPEG streams using inter-coding. That is, Index Tables provide a mechanism to efficiently locate stored Picture information based in display order.

The same mechanism shall be followed by MXF decoders receiving MXF files with VC-1 encoded bitstreams. For the purpose of Index Table temporal re-ordering, the concept of I-picture, P-picture and B-picture used in SMPTE 377-1, can be considered to have the same meaning in both MPEG and VC-1 coding schemes.

12.1 Indexing BI-pictures and Skipped-pictures

BI-pictures are B-pictures that include only intra-coded information. Following the above, BI-pictures shall be indexed using the same mechanisms used for B-pictures.

The “Flags” attribute of the Index Entry associated with a BI-picture (according to the Structure of Index Entry Array table in SMPTE 377-1) shall have bits 4 and 5 set to ‘0’ (zero) indicating there is no inter-coding in this picture. Nevertheless, the property “Key-Frame Offset” in that same table, shall be set according to the number of edit units to the previous I-picture. This enables MXF decoders to analyze GOP structure at ay point in the stream using Index Table constructs, clearly differentiating BI-pictures from I-pictures.

Skipped-pictures are P-pictures which are identical to their reference pictures. Following the above, skipped-pictures shall be indexed using the same mechanisms used for P-pictures.

Annex A (Informative)
Bibliography

ANSI/SMPTE 298M-1997, Television — Universal Labels for Unique Identification of Digital Data

SMPTE 336M-2007, Data Coding Protocol Using Key-Length-Value

SMPTE RP 210, Metadata Dictionary Registry of Metadata Element Descriptions

SMPTE RP 224, SMPTE Labels Registry

SMPTE EG 41-2004, Material Exchange Format (MXF) — Engineering Guideline