

SMPTE STANDARD

Material Exchange Format (MXF) — Mapping MPEG Streams into the MXF Constrained Generic Container



Page 1 of 39 pages

Table of Contents	Page
Foreword	2
Intellectual Property	3
Introduction.....	3
1 Scope	4
2 Conformance Notation	4
3 Normative References	4
4 Glossary of Acronyms, Terms and Data Types	5
5 MPEG Wrapping	6
5.1 Basics of Wrapping: Frames and Access Units	6
5.2 Wrapping Options	6
5.2.1 Overview	6
5.2.2 Empty and Variable Elements	6
5.2.3 MPEG in the Generic Container	7
5.2.4 Interleaving Guidelines	7
6 MXF Tracks	7
6.1 MXF Tracks and MPEG Content.....	7
7 Mapping MPEG into the Generic Container	8
7.1 Frame (Video AU) Wrapping	8
7.1.1 Interleaving Rules	9
7.1.2 Wrapping 3:2 Sequences	10
7.2 Clip Wrapping	10
7.3 Custom Wrapping	10
8 KLV Coding of MPEG Elements	11
8.1 KLV Coding for MPEG Video Streams	11
8.1.1 Picture Element Key	11
8.2 KLV Coding for MPEG Audio Streams	12
8.2.1 Sound Element Key	12
8.2.2 Multi-Channel Audio	13
8.3 KLV Coding for MPEG Streams Other than Video or Audio	13
8.3.1 Data Element Key	13
8.4 Use of KLV Fill Items	14
8.5 Use of KAG	14

9	SMPTE Label for Essence Container Identification	14
9.1	MPEG Essence Container Label Definition.....	14
10	Essence Descriptors for MPEG Mappings	16
10.1	MPEG Descriptors	16
10.2	MPEG Video Descriptor	16
10.2.1	Key for the MPEG Video Descriptor	18
10.3	MPEG-4 Visual Sub Descriptor	18
10.3.1	Key for the MPEG-4 Visual Sub Descriptor	20
10.3.2	Class model of the MPEG-4 Visual Sub Descriptor (Informative)	20
10.4	MPEG Audio Descriptor	21
10.4.1	Key for the MPEG Audio Descriptor	21
11	Index Table Requirements for MPEG Mappings	21
11.1	Setting the Properties in an Index Entry	21
Annex A	Bibliography (Informative)	24
Annex B	Index Tables for MPEG Mappings (Informative).....	25
B.1	Edit Units and Index Tables	25
B.2	Index Table Example for Frame Wrapped CBE MPEG Essence	25
B.3	Index Table Example for Clip Wrapped CBE MPEG Essence	27
B.4	Index Table Example for Frame Wrapped VBE MPEG Essence with Extra Items in the Generic Container	30
B.5	Index Table Example for Clip Wrapped VBE MPEG Essence with Extra Items in the Generic Container	35
B.6	Indexing of Fill Items	38
Annex C	Identifying MPEG Picture Type and Finding MPEG Content on the Timeline (Informative) ...	39
C.1	Identification of MPEG Picture Type	39
C.1.1	Identifying MPEG Picture Type Using Frame Wrapping Without an Index Table	39
C.1.2	Identifying MPEG Picture Type Using an Index Table	39

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 381-2 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.

Introduction

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

SMPTE ST 381 is a multi-part standard.

SMPTE ST 381-1 is the renumbered archived version of SMPTE 381M-2005. SMPTE ST 381-2, this part, has been created to update SMPTE 381M-2005 by providing clarification of text and updating existing provisions to promote interoperability. Users of SMPTE ST 381-1 and SMPTE ST 381-2 need to be aware that SMPTE ST 381-1 refers only to the Generic Container within SMPTE ST 379-1, not SMPTE ST 379-2, while SMPTE ST 381-2 refers only to the Generic Container within SMPTE ST 379-2, not SMPTE ST 379-1.

SMPTE ST 381-2, this part, defines the synchronization requirements and mechanisms for mapping MPEG material within MXF including the definition of the MPEG Video and Audio Descriptors, Essence Element Keys, and an Essence Container Label for MPEG streams into MXF along with CBE and VBE indexing of MPEG content using Index Tables.

The SMPTE ST 381 multi-part specification on the MXF MPEG mapping was developed using a number of fundamental requirements. These requirements are the following:

- Preserve timing in any MPEG system stream.
- Preserve timing when de-multiplexing a system stream and converting to MPEG ES wrapping.
- Preserve the MPEG buffer model so that when the MPEG is removed, the buffer model is preserved when this is required.
- Encapsulate MPEG PES without knowledge of the contents.
- Provide a mechanism for key frame identification.
- No change of original MPEG data order, this will facilitate recording an MPEG stream in MXF and playing out an MPEG stream from MXF.
- Small CPU and memory needs for both recording and playing out.
- Covers all possible types of MPEG data: MPEG-2 TS, PS, MPEG-1 system, MPEG-1 and MPEG-2 PES, video ES, audio ES, private streams and other derived format such as DVD-VOB (DVD Video Object), VCD (Video CD), etc.
- Facilitate access to individual Access Units.
- Possibility to wrap the MPEG stream based on main level elements for storage critical environments.

1 Scope

This standard specifies the mapping of video and audio MPEG ES, as identified by an ISO/IEC 13818-1 stream_id value, as well as the carriage of data streams extracted from higher level MPEG multiplexes, into the MXF Generic Container described in SMPTE ST 379-2. This includes, but is not limited to, ISO/IEC 13818-2 MPEG video stream, ISO/IEC 11172-2 MPEG video streams, ISO/IEC 14496-2 MPEG visual streams, ISO/IEC 14496-10 video streams, ISO/IEC 13818-3 MPEG audio streams, ISO/IEC 11172-3 MPEG audio streams, and ISO/IEC 14496-3 MPEG audio streams. It does not include ISO/IEC 14496-15 video streams. The document outlines single stream mapping and synchronized multi-stream mapping along with synchronization requirements for the MXF encapsulation of co-timed MPEG Essence streams. This standard defines SMPTE universal labels to be used to uniquely identify specific MPEG implementations. The Essence Element Keys and an Essence Container Label for MPEG streams are detailed along with CBE and VBE indexing of MPEG content using Index Tables.

In order to achieve interoperability within any given Operational Pattern, restrictions may be placed on the way in which this Essence Container 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 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. 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 336:2007, Data Encoding Protocol using Key-Length-Value

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

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

ISO/IEC 13818-1:2007, Information Technology — Generic Coding of Moving Pictures and Associated Audio Information — Part 1: Systems

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

ISO/IEC 13818-3:1998, Information Technology — Generic Coding of Moving Pictures and Associated Audio Information — Part 3: Audio

ISO/IEC 13818-7:2004, Information Technology — Generic Coding of Moving Pictures and Associated Audio Information — Part 7: Advanced Audio Coding (AAC)

ISO/IEC 14496-2:2004, Information Technology — Coding of Audio Visual Objects — Part 2: Visual

ISO/IEC 14496-10:2009, 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 Glossary of Acronyms, Terms and Data Types

For the purposes of this document, the terms and definitions given in the previous section containing the normative references within this document and the following apply.

AAC: Advanced Audio Coding

Audio AU: The full definition of an Access Unit (AU) for audio is given in ISO/IEC 13818-1, but is summarized as the following: 1 Audio AU = The coded representation of an audio frame.

AVC/H.264: Advanced Video Coding – ISO/IEC 14496-10 (Part 10) | ITU-T Rec. H.264

BWF: Broadcast Wave Format for audio (see SMPTE ST 382)

ES: Elementary Stream

GOP: Group of Pictures

GOV: Group of VOP

MPEG-1: ISO/IEC 11172

MPEG-2: ISO/IEC 13818

MPEG-4 Visual: ISO/IEC 14496-2

Video AU: The full definition of an Access Unit (AU) for video is given in ISO/IEC 13818-1, but is summarized as the following: 1 Video AU = The coded data for a picture and any stuffing that follows it up to, but not including, the start of the next AU. If a picture is *not* preceded by a group_start_code or a sequence_header_code, the AU begins with a picture_start_code. If a picture is preceded by a group_start_code or a sequence_header_code, the AU begins with the first byte of these start codes. If it is the last picture preceding a sequence_end_code in the bit stream all bytes between the last bytes of the coded picture and the sequence_end_code (including the sequence_end_code) belong to the AU.

VO: Visual Object

VOL: Video Object Layer

VOP: Video Object Plane

VOS: Visual Object Sequence

5 MPEG Wrapping

5.1 Basics of Wrapping: Frames and Access Units

The MPEG video specifications enable the compression of picture-based video. These pictures can be frames or fields. The MPEG specification provides signaling for the number of pixels in the picture, the field/frame rate and other parameters including the aspect ratio of the pixels. Many of these parameters will be copied into the MXF Essence Descriptor.

The text in this document refers to Access Units (AUs) rather than pictures. The full definition of an AU for video and audio is given in ISO/IEC 13818-1 and a summary of an Audio AU and a Video AU is given in the glossary of this document.

Note: The Video AU Duration might not be the same as Audio AU Duration in an MPEG stream.

5.2 Wrapping Options

5.2.1 Overview

The MXF Generic Container described in SMPTE ST 379-2 consists of one or more contiguous KLV wrapped Content Packages. When wrapping every Video AU of long GOP MPEG, the frames shall be wrapped in transmission order rather than display order. When interleaving uncompressed audio samples in the same Content Package, the audio samples shall not be reordered to match the video. If the video is at a frame rate of 29.97 fps and the audio is sampled at 48 kHz, then there will be a varying number of audio samples in the Content Packages as a result of the complex relationship of the clocks.

5.2.2 Empty and Variable Elements

When a wrapping algorithm does not require data from a particular element, the element may be empty within the Content Package.

Elements within a Content Package may vary in size from one Content Package to another. These variations can be quite large and in the limit, the data for an element may not be present at all. In the case where a particular element has no Essence data to wrap, the KLV coded element shall be empty within the Content Package (i.e. the length will be zero and there will be no value data)..

SMPTE ST 379-2 contains further details on requirements for elements within a Content Package.

5.2.3 MPEG in the Generic Container

Figure 1 shows an example of a Generic Container, defined in SMPTE ST 379-2, to contain an MPEG stream. In general, the Generic Container was designed to create a streamable format where each repetition of the interleaved elements represents a single field or frame. Figure 1 shows an example that has one field of picture in every Content Package which is associated with an optional system item, some audio information and some data. This follows the Generic Container policy that the audio and data are co-timed with the video information in the same Content Package.

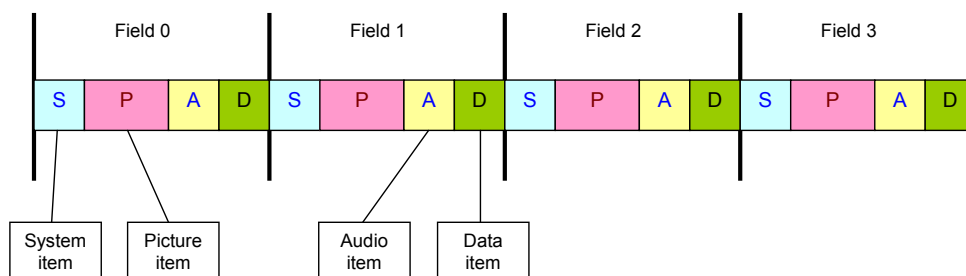


Figure 1 – Example of MPEG in a Generic Container

In some cases, different Content Packages will contain different Durations. When the compressed audio item is not synchronized with the picture item, the offset in time from the start of the audio Content Item to the Picture Content Item should be expressed using the fractional position offset mechanism defined in SMPTE ST 377-1 if an Index Table is constructed.

Note: Index Table requirements for MPEG mappings are discussed in Section 11 of this document.

As stated previously, SMPTE ST 379-2 contains further details on requirements for elements within a Content Package and the use of a Generic Container.

5.2.4 Interleaving Guidelines

In order to give MXF encoder guidance for creating consistent Generic Containers between applications, interleaving guidelines are defined within this document. The principle of creating the interleaving guidelines is to allow the creation of MXF files containing a mixture of unspecified components. For example:

Long GOP MPEG ES + BWF sound;
Uncompressed pictures + long GOP MPEG proxy + AES sound.

Different mapping mechanisms are listed in this and other mapping documents. Specific interleaving guidelines for implementing those mapping schemes are provided to improve interoperability. In addition to these low level interleaving guidelines, there may be higher level Partition multiplexing rules operating which may be determined by the Operational Pattern of the file.

6 MXF Tracks

6.1 MXF Tracks and MPEG Content

Many MXF files containing MPEG content will contain a picture Track for the MPEG video Essence and sound Tracks for any audio Essence. Many professional applications will use uncompressed audio which will have a simple timing relationship with the MPEG video.

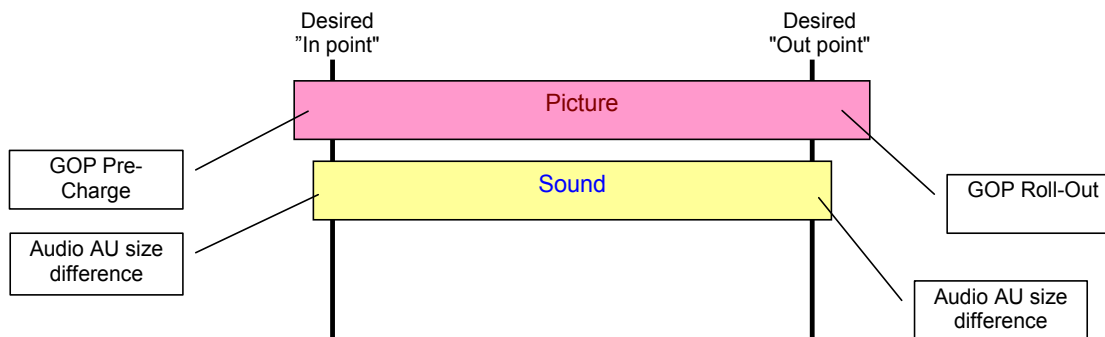


Figure 2 – Example of MXF Track Timing for MPEG Content

The example shown in Figure 2 shows the relationship between the desired start and end points of a stored generalized MPEG segment and the actual data start and end points which would be found in the stored Essence. The Track and sequence items within the header are used to accurately describe the synchronization scenario above.

As shown in Figure 2, the Pre-Charge used for inter-coded material represents the previous Edit Units that are needed before desired start Edit Unit of the Essence in order to correctly display the desired start picture. To display the last picture described by the Duration of the associated Essence Track, Roll-Out consisting of the count of Edit Units minus Origin minus the number of displayed pictures in the Essence Container may be needed for inter-coded material. Roll-Out is used to preserve the correct temporal order when displaying the last picture within the Essence. Taking Pre-Charge and Roll-Out into account, the value of the Container Duration Property within the File Descriptor, as described in SMPTE ST 377-1, equals the count of Edit Units starting at the first Content Package with a non-empty Essence Element to the last Content Package with a non-empty Essence Element.

In an OP1a file, the Material Package has the same Duration as the Top-Level File Package. The Duration Property of the Sequence Set for the Top-Level File Package consists of Pre-Charge and the Package Duration, but does not include Roll-Out. Using Pre-Charge and Roll-Out as described previously, unwanted frames, which were captured to allow decoding of the desired frames, are hidden.

Note: The audio can also be located in an external file. When the audio is external, there will be at least two File Packages.

If time code is present in the MPEG GOP headers or has been inserted in accordance with SMPTE ST 328 then, in the absence of specific application requirements, this time code should be used by the Material Package Timecode Track.

Note: Examples of an application specification is EBU R122.

7 Mapping MPEG into the Generic Container

7.1 Frame (Video AU) Wrapping

A sequence of images shall be KLV coded as defined in SMPTE ST 336. One type of KLV wrapping, defined as Frame Wrapping, is specified for the MPEG body structure and described in this section.

As described in SMPTE ST 379-2, Frame Wrapping has one or more Content Packages in the Essence Container. Examples of "Frame Wrapping" methods for MPEG are shown in Figures 3 and 4. Figure 3 shows an example of an MPEG video stream wrapped in a number of Content Packages with no other Generic Container elements in the container. Each Content Package has the Duration of one MPEG Video AU. Specifying the Duration through a MPEG Video AU determines where the MPEG headers will be found. Through these MPEG headers, the picture type can be determined as shown in Annex C of this document.

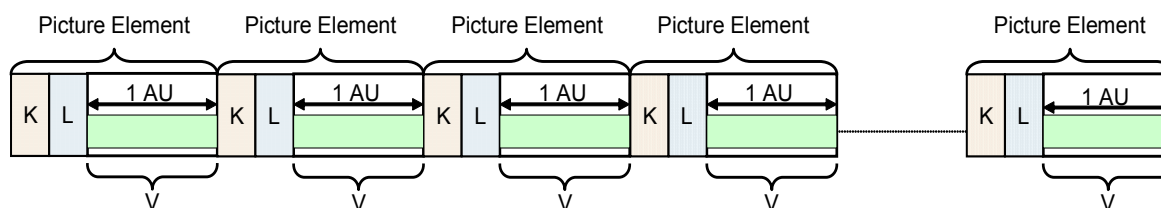


Figure 3 – Example of Frame Wrapping for MPEG Picture Elements

The Frame Wrapping method is intended to enable retrieval of individual AUs and the corresponding field/frame within them by MXF applications which process at the KLV level. This can be particularly useful for applications which support multiple Generic Container mapping types. Each MPEG Video AU is KLV wrapped using a Generic Container Picture Element Key. Using the Generic Container Picture Element Key, an individual MPEG Video AU within a KLV packet can be located as previously mentioned.

In some applications, the Frame Wrapped MPEG picture data can exist in Content Packages with other elements such as sound and data elements. An example Generic Container is shown in Figure 4.

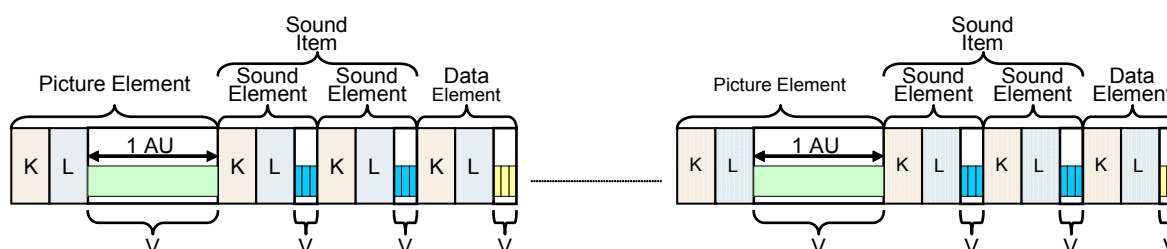


Figure 4 – Example of Frame Wrapping of MPEG Picture Elements with other Generic Container Elements

The Generic Container mapping specifications for the sound and data elements will detail the key values and format of the data within the elements. There are many situations where each wrapped frame of MPEG picture data will not exactly coincide with the sound/data elements within the same Content Package. These situations are described in more detail within SMPTE ST 379-2.

7.1.1 Interleaving Rules

When creating an MXF file from raw ES components, an MXF encoder has some flexibility. Easy access can be provided to all of the components by creating an MXF file for interchange. In this case, it is desirable to group together elements which are synchronized on the timeline. The overall goal is to keep the synchronized video, audio and associated data in the same Content Package. SMPTE ST 379-2 contains further details on requirements for elements within a Content Package and the use of a Generic Container.

7.1.2 Wrapping 3:2 Sequences

The MPEG coding of 3:2 sequences is defined in the MPEG specifications as explained in ISO/IEC 13818-2. KLV wrapping of the 3:2 AUs for interchange is, therefore, defined in the previous section. Ensuring that an MPEG Video AU is wrapped defines the KLV wrapping point regardless of the regularity of the 3:2 sequence.

Construction of Index Tables for 3:2 sequences requires the Index Table construction application to calculate the correct byte offset for a given temporal offset regardless of the MPEG encoding used. This may be done by the inspection of the MPEG headers for the occurrence of repeat_first_field flags or by some equivalent method.

7.2 Clip Wrapping

A sequence of images shall be KLV coded as defined in SMPTE ST 336. One type of KLV wrapping, defined as Clip Wrapping, is specified for the MPEG body structure and described in this section.

As described in SMPTE ST 379-2, Clip Wrapping is where the entire Duration of the Essence Container is contained with a single Content Package. An example of “Clip Wrapping” method for an MPEG Picture Element is shown in Figure 5. Figure 5 shows an MPEG video stream wrapped in a Content Package with no other Generic Container elements in the container.

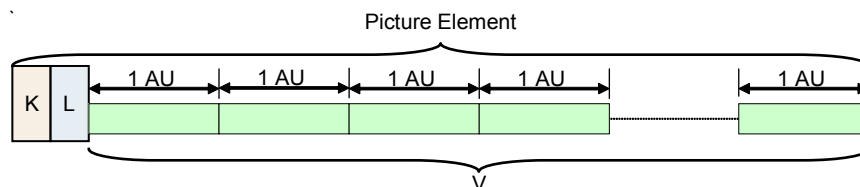


Figure 5 – Example of Clip Wrapping for an MPEG Picture Element

The Clip Wrapping method is intended for applications which carry the MPEG stream as a single large entity. This can be very useful in applications such as store and forward servers which process whole files and also in applications where it is desired to use the rich metadata structures of MXF as an annotation to MPEG data. The clip of MPEG data is KLV wrapped using an appropriate key as detailed in the KLV Coding for MPEG Streams other than Video or Audio section of this specification. When MPEG data is clip wrapped, there shall be only one clip per Generic Container. Multiple clips can be concatenated and edited using higher Operational Patterns detailed in SMPTE ST 377-1.

7.3 Custom Wrapping

A sequence of images shall be KLV coded as defined in SMPTE ST 336. One type of KLV wrapping, defined as Custom Wrapping, is specified for the MPEG body structure and described in this section.

In Custom Wrapping, KLV encoding wraps the MPEG data in an application specific way. All new wrapping modes shall be defined in a separate SMPTE standard and the associated key and label values registered in the SMPTE registry. Unless otherwise stated, each of the Custom Wrappings shall treat the MPEG video stream as the master for determining the start and end points of the Content Package. SMPTE ST 379-2 contains further details on requirements for elements within a Content Package and the use of a Generic Container.

A table of pre-defined Custom Wrapping modes shall be as defined in Table 1.

Table 1 – Custom wrapping descriptions

Name	Constraint	Application
Stripe	One KLV per macroblock row	Each Content Package contains an entire row of macroblocks. For some MPEG-1 video streams this mode may not be possible.
Fixed Audio Size	Fixed Audio Samples	The Sound Element is the Master Element. Wrapping is determined by a fixed number of audio samples. This can be used when the video rate is $1000/1001 * \text{Integer_rate}$. The Picture Elements are fitted into the closest Content Package to their start time.
Splice	KLV at splice points	Each Content Package starts and ends at a valid MPEG splice point.
Closed GOP	KLV at closed GOPs	Each Content Package starts and ends at a closed GOP.
Slave	Slaved to another constraint	The audio is slave wrapped. For example a video elementary stream can be wrapped in chunks whose size corresponds to the original PES packet payload, and the associated audio is wrapped in chunks of approximately the same Duration as the video.
No specific constraint		No specific wrapping constraints were employed.

Note: Byte 16 of the MPEG Essence Container Label specified in Table 6 of this document describes the signalling for the specific custom wrapping listed in Table 1.

8 KLV Coding of MPEG Elements

8.1 KLV Coding for MPEG Video Streams

When using this specification, all MPEG video streams shall be wrapped using a Picture Element Key. All MPEG audio streams shall be wrapped using a Sound Element Key. All other MPEG streams shall be wrapped using a Data Element Key.

8.1.1 Picture Element Key

The values of the first 12 bytes of the Picture Element Key shall be as defined in SMPTE ST 379-2. The values of the last four bytes of the Picture Element Key shall be as defined in Table 2:

Table 2 – Key Value for the MPEG Picture Element

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

8.1.1.1 Essence Element Count – Byte 14

This shall be the count of the number of Picture Elements in the Picture Item of the Content Package.

8.1.1.2 Essence Element Type – Byte 15

For Frame Wrapped MPEG video AUs, this shall be 05_h. For Clip Wrapped MPEG video AUs, this shall be 06_h. For Custom Wrapped MPEG video AUs, this shall be 07_h.

8.1.1.3 Essence Element Number – Byte 16

This is a number used as an index to identify this instance of the element type within the item in each Content Package. The Generic Container specification requires that this byte be unique within an item. Each instance of the MPEG Picture Element shall have a constant Essence Element number throughout the entire Generic Container.

Note: The actual value of this number has no meaning; however, its uniqueness is important.

8.1.1.4 Picture Element Length

The length field of the KLV coded element shall be 4 bytes BER long-form encoded (i.e., 83_h.xx.yy.zz) for Frame Wrapping. The length field of the KLV coded element shall be 8 bytes BER long-form encoded (i.e., 87_h.aa.bb.cc.dd.ee.ff.gg) for Clip Wrapping. For Custom Wrapping, the length field shall be consistent and shall be either 4 bytes or 8 bytes as appropriate.

8.1.1.5 Picture Element Value

For Frame Wrapping, the value field of the KLV coded element shall be one Video AU or empty, having a Duration of one Video AU even if it is empty.. For Clip Wrapping, this shall be the entire MPEG video clip where the Duration shall be two or more Video AUs. For Custom Wrapping, the video will be wrapped in an application specific way, and the wrapping shall be defined in a separate SMPTE standard and the associated key and label values registered in the SMPTE registry.

8.2 KLV Coding for MPEG Audio Streams

8.2.1 Sound Element Key

The values of the first 12 bytes of the Sound Element Key shall be as defined in SMPTE ST 379-2. The values of the last four bytes of the Sound Element Key shall be as defined in Table 3.

Table 3 – Key Value for the MPEG Sound Element

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

8.2.1.1 Essence Element Count – Byte 14

This shall be the count of the number of Sound Elements in the Sound Item of the Content Package.