

SMPTE STANDARD

Interoperable Master Format —
Core Constraints



Table of Contents	Page
Foreword	2
Intellectual Property	2
Introduction	2
1 Scope.....	3
2 Conformance Notation.....	3
3 Normative References.....	3
4 Overall	4
5 Track Files	5
6 Composition.....	14
7 IMF Master Package (IMP).....	18
Annex A Bibliography (Informative).....	20
Annex B Consolidated Schema (Informative)	21
Annex C Data Essence Definitions.....	22
Annex D StereolImageTrackFileResourceType (Normative).....	23
Annex E Sparse Data Essence (Normative)	24
Annex F Additional Generic Sound Essence Descriptor Items (Normative).....	25
Annex G Alternative Center Cuts (Normative).....	26
Annex H Active Area Rectangle (Normative)	28
Annex I RFC 5646 Language Tag List (Normative).....	31

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 2067-2 was prepared by Technology Committee 35PM.

Intellectual Property

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

Introduction

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

A single TV or movie title is transformed into multiple content versions (airline edits, special edition, languages...). These versions, which share common assets sourced from high-quality source masters, are ultimately made available to multiple distribution channels (Internet, optical media, broadcast...) across multiple territories and over the span of many months to over a year.

The IMF is a file-based framework that allows these high-quality versions, called Compositions, to be efficiently represented, managed, played back, processed and transformed on file-based systems. For example, it facilitates the generation of multiple outputs of the same Composition (through instructions contained in an Output Profile List define in other documents) to accommodate the specific needs of distribution channels. Since management and processing of Compositions are performed across multiple devices and service providers, interoperability is desirable.

1 Scope

This document specifies provisions that may be used by multiple IMF Applications. These provisions include essence and metadata constraints, but exclude image essence constraints, which are specified along with other parameters by each IMF 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 382:2007) although, during a transitional phase, the document as published (printed or PDF) may bear an older designation (such as SMPTE 382M-2007). Documents with the same root number (e.g. 382) and publication year (e.g. 2007) 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 RP 224, SMPTE Labels Registry

SMPTE ST 330:2011, Television — Unique Material Identifier (UMID)

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

Amendment 1:2012 to SMPTE ST 377-1:2011

Amendment 2:2012 to SMPTE ST 377-1:2011

SMPTE ST 377-4:2011 Material Exchange Format (MXF) - MXF Multichannel Audio Labeling Framework

SMPTE ST 382:2007 Material Exchange Format — Mapping AES3 and Broadcast Wave Audio into the MXF Generic Container

Amendment 1:2012 to SMPTE ST 382:2007

Amendment 2:2013 to SMPTE ST 382:2007

SMPTE ST 429-5:2009 D-Cinema Packaging — Timed Text Track File

SMPTE ST 429-8:2007 D-Cinema Packaging — Packing List

SMPTE ST 430-2:2006 D-Cinema Operations — Digital Certificate

SMPTE ST 436-1:2013, MXF Mappings for VI Lines and Ancillary Data Packets

SMPTE ST 2001-1:2013, XML Representation of SMPTE Registered Data — Mapping Rules

SMPTE ST 2029:2009 Uniform Resource Names for SMPTE Resources.

SMPTE ST 2052-1:2010, Timed Text Format (SMPTE-TT)

SMPTE RP 2052-10:2012, Conversion from CEA-608 Data to SMPTE-TT

SMPTE ST 2067-3:2013, Interoperable Master Format – Composition Playlist

SMPTE ST 2067-5:2013, Interoperable Master Format – Essence Component

SMPTE ST 2067-8:2013, Interoperable Master Format – Common Audio Channels, Soundfield Groups and Groups of Soundfield Groups.

ISO/IEC 14496-18:2004. Font Compression and Streaming

Internet Engineering Task Force (IETF) (November 1996). RFC 2046 Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types.

Internet Engineering Task Force (IETF) (2005, April). RFC 4051 – Additional XML Security Uniform Resource Identifiers (URIs).

Internet Engineering Task Force (IETF) (July 2005). RFC 4122 A Universally Unique Identifier (UUID) URN Namespace.

Internet Engineering Task Force (IETF) (2009, September). RFC 5646 – Tags for Identifying Languages.

Internet Engineering Task Force (IETF) (2006, May). RFC 5939 – Media Type Registration for the Society of Motion Picture and Television Engineers (SMPTE) Material Exchange Format (MXF).

World Wide Web Consortium (W3C) (2010 November 18). Timed Text Markup Language (TTML) 1.0

World Wide Web Consortium (W3C) (2004 October 28). XML Schema Part 1: Structures (Second Edition).

World Wide Web Consortium (W3C) (2004 February 4). Extensible Markup Language (XML) 1.0 (Third Edition).

World Wide Web Consortium (W3C) (2002 February 12). XML Signature Syntax and Processing.

4 Overall

4.1 XML Schema and Namespace

XML elements defined by this specification shall conform to the XML schema definitions (see W3C XML Schema Part 1: Structures) found in this specification. In the event of a conflict between schema definitions and the prose, the prose shall take precedence.

The XML schema root element shall be as defined in Table 1.

Table 1 – XML Schema root element definition

```
<xs:schema targetNamespace="http://www.smpte-ra.org/schemas/2067-2/2013"
            xmlns:cpl="http://www.smpte-ra.org/schemas/2067-3/2013"
            xmlns:cc="http://www.smpte-ra.org/schemas/2067-2/2013"
            xmlns:xs="http://www.w3.org/2001/XMLSchema"
            elementFormDefault="qualified" attributeFormDefault="unqualified">
  <xs:import namespace="http://www.smpte-ra.org/schemas/2067-3/2013" />
  <!-- schema definitions found in this document -->
</xs:schema>
```

4.2 UUID Generation

UUID values used by this specification shall be generated as specified in IETF RFC 4122. UUID values that identify assets or cryptographic information shall be generated using a truly-random or pseudo-random number source, and shall have a Version field value of '4' (or 0100b).

Note: The 'b' suffix on this value indicates a binary encoding, most significant bit (MSB) first.

4.3 XML Character Encoding

XML documents shall use UTF-8 encoding, as specified in W3C Extensible Markup Language (XML).

5 Track Files

5.1 Overall

5.1.1 Format

Track Files shall conform to SMPTE ST 2067-5.

5.1.2 Shim Parameters

Track Files shall be associated with the shim parameter values (see SMPTE ST 2067-5) specified in Table 2.

Table 2 – Shim Parameter Values Definitions

<i>Shim Parameter</i>	<i>Value</i>	<i>Notes</i>
audio_family	ST 2067-2	See Section 5.3.2
audio_file_arrangement	ST 382	See Section 5.3.3
data_family	ST 436-1, ST 2052-1	See Sections 5.4.3 and 5.5
data_file_arrangement	ST 436-1, ST 429-5	See Section 5.4.2 and 5.5
sys_item_tc	true	Metadata in the System Item can be copied or moved to header metadata.
partition_spacing	60s	See Section 5.1.3
index_strategy_frame	Follow	See Section 5.1.4
Index_strategy_clip	Follow	See Section 5.1.4
track_tag_policy	none	See Section 5.3.6
ingest_TC	none	No dominant source of timecode during ingest is specified.
lead_TC	CPL	The Composition Playlist (CPL) specified in ST 2067-3 defines a synthetic timecode Composition timeline.
dropframe_tc	none	No constraints are specified on timecode tracks in Track File header metadata.

Note: The Composition Playlist specified in SMPTE ST 2067-3 provides synchronization across the Composition timeline and timecode is not used.

Table 3 – Shim Parameter defined by Applications

<i>Shim Parameter</i>
shim_id
gc_type
picture_family
picture_bitrate
picture_format
picture_custom_ANC
picture_render_ANC

The shim parameters listed in Table 3 shall be defined by individual Applications.

5.1.3 Body Partition

With the exception of the last body partition, the body partitions of a frame-wrapped Track File shall occur every 60 ± 6 s.

5.1.4 Index Tables

At each partition point in a Track File, the index partition shall follow the essence partition that it indexes.

5.1.5 Identification

The Package UID of a given Track File shall be a basic UMID per SMPTE ST 330, having a UUID value in the material number part and a material number generation method of UUID/UL. The Package UID value shall be further constrained as follows:

- Byte 11 of the UL portion of the UMID shall be 0Fh (unidentified material type).
- Byte 12 of the UL portion of the UMID shall be 20h (UUID/UL material number generation method and undefined instance number generation method).
- The three bytes of the instance number shall be 0 (zero).

Package UID values generated in accordance with the normative provisions of this subsection will thus have the following contents in the first 16 bytes: 060a2b34h 01010105h 01010f20h 13000000h.

5.1.6 MIME Type

The media type of a Track File shall be application/mxf as specified in IETF RFC 4539.

5.2 Image Track Files

5.2.1 General

An Image Track File shall conform to Sections 5.1 and 5.2.

An Image Track File primarily contains image essence. This specification does not specify the nature of the image essence.

5.2.2 Alternative Center Cut

The item defined in Annex G may be present in the Generic Picture Essence Descriptor (see SMPTE ST 377-1).

5.2.3 Active Area Rectangle

The items defined in Annex H may be present in the Generic Picture Essence Descriptor (see SMPTE ST 377-1).

5.3 Audio Track Files

5.3.1 General

An Audio Track File shall conform to Sections 5.1 and 5.3.

An Audio Track File primarily contains audio essence.

Each Audio Track File is intended to carry a single soundfield group. If the audio that is to be carried consists of more than one soundfield group, then each is carried in its own Audio Track File(s). For example, 6+2 Music and Effects content consists of 3 Audio Track Files: one containing the "Music and Effects Main" 5.1 Soundfield Group and two others containing each a single center channel, one for "Music and Effects Optional" and one for "Dialog Guide". Another example is, 6+2 Music and Effects would be carried by 8 Audio Track Files: 6 containing the "Music and Effects Main" 5.1 Soundfield Group and two others containing each a single center channel, one for "Music and Effects Optional" and one for "Dialog Guide".

5.3.2 Essence

5.3.2.1 General

The audio essence shall consist of an audio waveform digitally sampled at regular interval into a sequence of audio samples.

5.3.2.2 Sampling Rate

The sampling rate of the audio samples shall be either 48000 Hz or 96000 Hz.

Note: The sampling rates specified above are absolute and specified independently of a video signal. The Composition Playlist (SMPTE ST 2067-3) provides the mechanism to synchronize audio and image essence.

5.3.2.3 Quantization

Each audio sample shall be uniformly quantized as a 24-bit signed integer in the range $[-(2^{23}) .. 2^{23} - 1]$.

Source audio uniformly quantized using signed integers with lower bit depth, e.g. 16- or 20-bit, shall be padded in the least significant bits to create 24-bit audio samples.

5.3.3 Wrapping

The audio essence within an Audio Track File shall be:

- clip-wrapped as a Wave Clip-Wrapped Element, as specified in SMPTE ST 382; and
- packed bit-by-bit as specified by AES31-2, as specified in Section 7.2.1 of SMPTE ST 382.

5.3.4 Wave Audio Essence Descriptor

5.3.4.1 General

The Top-Level File Package shall reference a Wave Audio Essence Descriptor as specified in SMPTE ST 382.

The constraints specified in Table 4 shall apply. Items labeled "shall be present" shall be present in the Wave Audio Essence Descriptor. Items labeled "shall be ignored" shall be ignored if present.

Table 4 – Wave Audio Essence Descriptor Constraints

<i>Item</i>	<i>Constraints</i>
Locked/Unlocked	Shall be ignored
Audio Ref Level	Shall be ignored
Electro-Spatial Formulation	Shall be ignored
Dial Norm	Shall be ignored
Channel Assignment	See Section 5.3.4.2

5.3.4.2 Channel Assignment

The Channel Assignment item shall be present and shall be equal to the UL specified in Table 5.

Table 5 – Specification of the Channel Assignment Label

<i>Byte No.</i>	<i>Description</i>	<i>Value (hex)</i>	<i>Meaning</i>
1-7	Registry Designator	See SMPTE ST 400	
8	Registry Version Number	0Dh	Version of SMPTE RP 224 in which this label first appears
9	Parametric	04h	Node used to define parametric data
10	Sound Essence	02h	Identifies sound essence coding
11	Sound Coding Characteristics	02h	Identifies sound coding characteristics
12	Sound Channel Labeling	10h	Identifies sound channel labeling
13	Sound Channel Labeling SMPTE ST 2067-2	04h	Identifies sound channel labeling as defined in SMPTE ST 2067-2
14	Application of the MXF Multichannel Audio Framework	01h	Indicates that the MXF Multichannel Audio Framework is used
15	Reserved	00h	Reserved
16	Reserved	00h	Reserved

5.3.5 Additional Generic Sound Essence Descriptor Items

The items specified in Annex F shall be subject to the constraints specified in Table 6.

Table 6 – Additional Generic Sound Essence Descriptor Items Constraints

<i>Item</i>	<i>Constraints</i>
Reference Audio Alignment Level	Should be present.
Reference Image Edit Rate	Shall be present and shall be equal to the Composition Edit Rate. See Section 6.4.

5.3.6 Multichannel Audio Labeling

5.3.6.1 General

An Audio Track File shall conform to SMPTE ST 377-4.

5.3.6.2 AudioChannelLabelSubDescriptor Instances

Each audio channel in an Audio Track File shall be associated with exactly one AudioChannelLabelSubDescriptor instance, and each AudioChannelLabelSubDescriptor instance shall be associated with exactly one audio channel.

5.3.6.3 SoundfieldGroupLabelSubDescriptor Instances

An Audio Track File shall contain exactly one SoundfieldGroupLabelSubDescriptor instance, which all AudioChannelLabelSubDescriptor instances shall reference.

5.3.6.4 GroupOfSoundfieldGroupsLabelSubDescriptor Instances

An Audio Track File shall contain zero or one GroupOfSoundfieldGroupsLabelSubDescriptor.

5.3.6.5 Item Constraints

Within a given Audio Track File, the constraints of Table 7 shall apply.

Table 7 – MCA SubDescriptor Constraints.

<i>Item</i>	<i>AudioChannelLabel SubDescriptor Constraints</i>	<i>SoundfieldGroup Label SubDescriptor Constraints</i>	<i>GroupOfSoundfieldGro ups Label SubDescriptor Constraints</i>
SoundfieldGroupLin kID	Shall be present. See Section 5.3.6.3.		
MCA Channel ID	If absent, a value of 1 shall be assumed. See Section 5.3.6.2.	Shall be ignored.	
MCA Tag Name	Shall be present.		
RFC 5646 Spoken Language	Shall be ignored.	Shall be equal to the primary spoken language associated with the soundfield group. It shall be absent if and only if the soundfield group is not associated with a primary spoken language.	Shall be ignored.
MCA Audio Content Kind	Shall be ignored.	Shall be present.	
MCA Element Kind			
MCA Title			
MCA Title Version			

Other items defined in SMPTE ST 377-4 but not required by this specification may be present and may be safely ignored by implementations.

Notes:

1 MCA Tag Symbol and MCA Tag Name contain a human-readable text intended for display to the user. The MCA Label Dictionary ID is used to unambiguously determine the nature of the underlying audio channel, soundfield group or group of soundfield groups.

2 MCA Audio Content Kind, MCA Element Kind, MCA Title and MCA Title Version contain human-readable descriptive text intended for display to the user. Grouping of audio channels and soundfield groups is achieved through SoundfieldGroupLabel SubDescriptor and GroupOfSoundfieldGroupsLabelSubDescriptor, respectively.

5.3.6.6 Common Audio Channels, Soundfield Group and Group of Soundfield Groups

The following applies for audio channels, soundfield group and group of soundfield groups identified using SMPTE ST 2067-8, which Audio Track Files should use whenever possible.

The MCA Label Dictionary ID, MCA Tag Symbol and MCA Tag Name item values shall be set according to Table 8.

Table 8 – Mapping of SMPTE ST 2067-8 parameters to SMPTE ST 377-4 item values

SMPTE ST377-4 item	SMPTE ST 2067-8 parameters
MCA Label Dictionary ID	UL parameter value
MCA Tag Symbol	Symbol prepended with the string <ul style="list-style-type: none"> • 'ch' for audio channels. • 'sg' for sound field groups. • 'gg' for groups of soundfield groups.
MCA Tag Name	Name

Not all audio channels belonging to a soundfield group need to be present in a given Audio Track File, but only AudioChannelLabelSubDescriptor instances associated with audio channels belonging to the soundfield group shall reference the SoundfieldGroupLabelSubDescriptor associated with the soundfield group.

The use of groups of soundfield groups is optional. A GroupOfSoundfieldGroupsLabelSubDescriptor instance may not be present even though some or all SoundfieldGroupLabelSubDescriptor instances associated with soundfield groups belonging to a given group of soundfield groups are present. However, only SoundfieldGroupLabelSubDescriptor instances associated with soundfield groups belonging to the group of soundfield groups may reference the GroupOfSoundfieldGroupsLabelSubDescriptor associated with that group of soundfield groups.

5.4 Data Essence Track Files

5.4.1 General

A Data Essence Track File shall conform to Sections 5.1 and 5.4.

A Data Essence Track File primarily contains data essence, e.g. subtitle and caption.

5.4.2 Wrapping

Data Essence Track Files shall conform to SMPTE ST 429-5.

The Timed Text Resource shall contain a single SMPTE-TT document, as specified in SMPTE ST 2052-1, which shall use UTF-8 character encoding and shall conform to the constraints specified in the following Sections.

5.4.3 Essence

Any of the generic font families defined in W3C Timed Text Markup Language may be referenced. Mapping to actual fonts is not specified.

The metadata element of the head element (see W3C Timed Text Markup Language) shall contain a single instance of the TimedTextResourceID element specified in Table 9. The UUID value of this element shall equal the UUID value of the Resource ID item of the TimedTextDescriptor.

Table 9 – TimedTextResourceID element schema definition

```
<xs:element name="TimedTextResourceID" type="dcm1:UUIDType"/>
```

Data Essence Track Files may contain data essence mapped from CEA-608 using SMPTE RP 2052-10, including a CEA-608 data tunnel.

If the data essence is associated with a single primary language, the xml:lang attribute of the root <tt> element shall identify the primary language.

If the data essence is associated with multiple languages, the xml:lang attribute of the root <tt> element shall be the empty string and the xml:lang attribute of each topmost element associated with a single language shall identify the language contained therein.

5.4.4 RFC 5646 Language Tag List

The Timed Text Descriptor may contain the RFC 5646 Language Tag List property specified in Annex I. The value of the property shall indicate the languages associated with the data essence within the Data Essence Track File. Absence of the property shall indicate that no language is associated with the Data Essence Track File.

5.4.5 Image Resources

For each distinct background image URI referenced in a Data Essence Track File, an image resource shall be carried as an Ancillary Resource in the same Data Essence Track File. The structure of each background image URI is not specified beyond the constraints specified in SMPTE ST 2052-1, but no two image resources referenced in a single Data Essence Track File shall have identical URI.

The format of each image resource shall be PNG as specified in SMPTE ST 2052-1.

The MIMEMediaType of each Ancillary Resource containing a font resource shall be "image/png".

The AncillaryResourceID of each Ancillary Resource containing an image resource shall be a Type 5 UUID as specified in IETF RFC 4122, using (i) background image URI as the name, (ii) SHA-1 as the algorithm, and (iii) the NameSpace_URL constant as the Name Space ID.

For instance, the URI "file:///animage.png" would result in an AncillaryResourceID with the value of 2d04c0ff-b414-5d42-8d99-5a0eff714b1a.

5.4.6 Font Resources

For each distinct font family that is (a) referenced in a Data Essence Track File and (b) not a generic font family as defined in W3C Timed Text Markup Language, a font resource shall be carried as an Ancillary Resource in the same Data Essence Track File

Each font resource shall conform to ISO/IEC 14496-18.

The MIMEMediaType of each Ancillary Resource containing a font resource shall be "application/x-font-opentype".

The AncillaryResourceID of each Ancillary Resource containing a font resource shall be a Type 5 UUID as specified in IETF RFC 4122, using (i) the font family name exactly as it appears in the data essence as the name, (ii) SHA-1 as the algorithm, and (iii) b6cc57a0-87e7-4e75-b1c3-3359f3ae8817 as the Name Space ID.

For instance, the URI "Arial" would result in an AncillaryResourceID with the value of fe16f8e4-57a9-56b5-a93c-6c27d6f61619.

5.5 Ancillary Data Track Files

An Ancillary Data Track File shall conform to Sections 5.1 and 5.5.

An Ancillary Data Track File shall conform to SMPTE ST 436-1.

An Ancillary Data Track File can be used to carry VBI and/or ANC data contained in source material. This specification neither constrains nor specifies this data.

6 Composition

6.1 Application Identification

The ExtensionProperties element of the Composition Playlist instance shall include a single instance of the ApplicationIdentification element specified in Table 10.

Table 10 – XML Schema definition for ApplicationIdentification element

```
<xs:element name="ApplicationIdentification">
  <xs:simpleType>
    <xs:restriction>
      <xs:simpleType>
        <xs:list itemType="xs:anyURI"/>
      </xs:simpleType>
      <xs:minLength value="1"/>
    </xs:restriction>
  </xs:simpleType>
</xs:element>
```

Each unique item of the ApplicationIdentification element shall identify an Application to which a Composition conforms.

Each Application shall define a value, appropriate for inclusion in ApplicationIdentification, that is uniquely associated with its normative provisions.

6.2 Homogenous Essence

6.2.1 Audio

Within a given a composition, the audio essence characteristics listed in Table 11 shall remain constant.

Table 11 – Homogeneous Audio Essence Characteristics

Characteristic	Definition
Quantization	See Section 5.3.2.3
Sampling Rate	See Section 5.3.2.2

6.3 Virtual Tracks

6.3.1 Main Image Virtual Track

A Composition shall contain exactly one Main Image Virtual Track.

The Main Image Virtual Track shall consist of one or more MainImageSequence elements, as specified in Table 12.

Table 12 – MainImageSequence element schema definition

```
<xs:element name="MainImageSequence" type="cpl:SequenceType"/>
```

If the underlying image essence consists of a sequence of pairs of image frames for stereoscopic viewing, a left eye frame and a right eye frame coincident in time, each MainImageSequence element shall contain Resource elements of type StereolImageTrackFileResourceType, as defined in Annex D, that reference two Image Track Files conforming to Section 5.2.

If the underlying image essence is monoscopic, all MainImageSequence elements shall contain Resource elements of type TrackFileResourceType that reference a Track File that conforms to Section 5.2.

The Edit Rate of the Resource elements shall be equal to the image frame rate of the underlying essence.

6.3.2 Audio Virtual Tracks

A Composition shall contain one or more Audio Virtual Tracks.

Each Audio Virtual Track shall consist of one or more MainAudioSequence elements, as specified in Table 13.

Table 13 – MainAudioSequence element schema definition

```
<xs:element name="MainAudioSequence" type="cpl:SequenceType"/>
```

All Audio Essence Track Files referenced by a given Virtual Track shall have identical sets of GroupOfSoundfieldGroupsLabelSubDescriptors, SoundfieldGroupLabelSubDescriptors and AudioChannelLabelSubDescriptors instances.

Each MainAudioSequence element shall contain Resource elements of type TrackFileResourceType that reference Audio Track File conforming to Section 5.3.

The Edit Rate of the Resource elements shall be equal to the audio sampling rate of the underlying essence.

6.3.3 Data Essence Virtual Tracks

A Composition shall contain zero or more Data Essence Virtual Tracks.

Each Data Essence Virtual Track shall consist of one or more instances of one of the element specified in Table 14.

Table 14 – Data Essence Sequence schema definition

```
<xs:element name="SubtitlesSequence" type="cpl:SequenceType"/>
<xs:element name="HearingImpairedCaptionsSequence" type="cpl:SequenceType"/>
<xs:element name="VisuallyImpairedTextSequence" type="cpl:SequenceType"/>
<xs:element name="CommentarySequence" type="cpl:SequenceType"/>
<xs:element name="KaraokeSequence" type="cpl:SequenceType"/>
```

Each element of Table 14 should contain Data Essence of the kind specified in Table 15.

Table 15 – Data Essence Element Sequence Content Kind

Data Sequence Element	Kind of Data Essence Content as defined in Annex C
SubtitlesSequence	Subtitle
HearingImpairedCaptionsSequence	Captions for the Hearing Impaired
VisuallyImpairedTextSequence	Text for the Visually Impaired
CommentarySequence	Commentary
KaraokeSequence	Karaoke

Each Resource elements within Data Essence Virtual Track:

- shall be of type TrackFileResourceType
- shall reference a Track File that conform to Section 5.4.
- shall have a native start point corresponding to time coordinate 0 in the timeline of the underlying Data Essence Track File.
- shall have a native duration equal to or larger than the time coordinate after which no element will be active in the timeline of the underlying Data Essence Track File.

Annex E provides guidance in the situation where there is no data essence associated with a particular Segment of the Composition.

6.3.4 Ancillary Data Virtual Track

A Composition shall contain zero or more Ancillary Data Virtual Tracks.

Each Ancillary Data Virtual Track shall consist of one or more instances of one of the element specified in Table 16.

Table 16 – Ancillary Data Sequence schema definition

```
<xs:element name="AncillaryDataSequence" type="cpl:SequenceType"/>
```

All Ancillary DataSequence elements shall contain Resource elements of type TrackFileResourceType that reference Track File that conform to Section 5.5.

6.4 Composition Edit Rate

The Composition Edit Rate shall be equal to the edit rate of the image essence referenced by the Main Image Virtual Track.

Note: The Composition Edit Rate does not constrain the rate at which implementations can reproduce a Composition. For instance, a Composition with an Edit Rate of 24/1 can be reproduced at 24,000/1,001 frames per second.

6.5 Track File Identification

The UUID value of the TrackFileId element of TrackFileResourceType instances shall be equal to the material number part of the Package UID of the Top-level File Package.

Per the constraints of Section 5.1.5, there is therefore a one-to-one mapping between the value of the TrackFileId element and the complete UMID value of the Package UID of the Top-level File Package of the Track File; and the UUID URN value of the TrackFieldId element shall be considered equivalent to the corresponding complete UMID URN, encoded as specified in Section 8.1 of SMPTE ST 2029.

For instance, a TrackFileId value of

urn:uuid:75b7e2a0-e5cc-11e1-aff1-0800200c9a66

is equivalent to the following UMID

urn:smp:umid:060a2b34.01010105.01010f20.13000000.75b7e2a0.e5cc11e1.aff10800.200c9a66

6.6 Segment Duration

The duration of a Segment shall be greater than or equal to the duration of one image essence frame referenced by the Main Image Virtual Track.

6.7 Content Version

A Composition Playlist instance shall contain at least one ContentVersion element.

The Id element within each ContentVersion element may contain any valid URI. A few schemes are listed in Table 17 for illustration.

Table 17 – Examples of URI schemes for Content Version identification (Informative)

Basic UMID (SMPTE ST 2029)
ISAN (IETF RFC 4246)
UUID (IETF RFC 4122)
Info (IETF RFC 4452)

6.8 EssenceDescriptor

The entire File Descriptor structure (including Sub Descriptors) referenced by the top-level File Package of each Track File referenced by the Composition Playlist shall be mapped to a single EssenceDescriptor element (see Section 7.1.10.1 in SMPTE ST 2067-3) using a single RegXML fragment as specified in SMPTE ST 2001-1.

In particular, instances of the following Descriptors and Sub Descriptors shall be mapped:

- Wave Audio Essence Descriptor (SMPTE ST 382) including the additional items defined in 0,
- MCA Label Descriptors (SMPTE ST 377-4),
- Timed Text Descriptor (SMPTE ST 429-5), and
- Timed Text Resource Descriptor (SMPTE ST 429-5).

6.9 Digital Signature

The Signature element of the Composition Playlist shall satisfy the following constraints:

- The KeyInfo element shall be present and shall contain the entire certificate chain for the signer.
- The Object element shall not be present and the URI attribute of the Reference element shall be the empty string, i.e. the signature is enveloped.
- The Reference element shall contain a single DigestMethod element, with its Algorithm attribute set to the URI value <http://www.w3.org/2001/04/xmlenc#sha256>.
- The Reference element shall contain a single Transform element, with its Algorithm attribute set to the URI value <http://www.w3.org/2000/09/xmldsig#enveloped-signature>.
- The CanonicalizationMethod shall be set to the URI value <http://www.w3.org/TR/2001/REC-xml-c14n-20010315>.
- The SignatureMethod shall be set to the URI value <http://www.w3.org/2001/04/xmldsig-more#rsa-sha256> as defined in IETF RFC 4051.

If X.509 certificates are used as specified in W3C XML Signature Syntax and Processing, then the entire certificate chain shall be carried in the KeyInfo element as a sequence of X509Data elements. Each of the X509Data elements shall correspond to one certificate in the chain, and contain one X509IssuerSerial element and one X509Certificate element. The Distinguished Name value in all X509IssuerName elements shall be compliant with RFC 2253 as specified in W3C XML Signature Syntax and Processing.

6.10 Digital Certificates

If the Signature element is present in a CPL instance, the digital certificate used shall conform to SMPTE ST 430-2.

7 IMF Master Package (IMP)

7.1 General

An Interoperable Master Package (IMP) shall consist of one Packing List, as specified in SMPTE ST 429-8, and all the assets it references.

A Packing List instances may reference assets which are referenced by other Packing List instances.

7.2 Asset Identification

The value of the Id element within each Asset element shall uniquely identify the asset and shall be unique in the Packing List instance.

The value of the Id element shall be extracted from the asset as specified in Table 18 for the asset defined therein.

Table 18 – Asset Identification

<i>Asset</i>	<i>Identifier</i>
Composition Playlist	Value of the Id element of the CompositionPlaylist instance.
Track File	Package UID value of the Top-level File Package.

The Packing List may reference assets not listed in Table 18.

7.3 Digital Signature and Certs

When the Signature element is present, digital certificates in the signer's certificate chain shall conform to the provisions of SMPTE ST 430-2.

7.4 Group ID

The GroupId element shall be ignored.

7.5 Complete IMP

A Complete IMP is an IMP containing only the complete set of assets comprising one or more compositions.

7.6 Partial IMP

A Partial IMP is an IMP containing one or more incomplete compositions (i.e., some assets needed to complete the composition are not present in the package.)

Annex A Bibliography (Informative)

SMPTE ST 2016-1:2009, Format for Active Format Description and Bar Data

Internet Engineering Task Force (IETF) (February 2006). RFC 4246 – International Standard Audiovisual Number (ISAN) URN Definition

Internet Engineering Task Force (IETF) (April 2006). RFC 4452 – The "info" URI Scheme for Information Assets with Identifiers in Public Namespaces

Annex B Consolidated Schema (Informative)

This document is accompanied by a file named *imf-core-constraints.xsd* that contains an informative schema instance combining all schema definitions specified herein.

Annex C Data Essence Definitions (Normative)

C.1 Subtitles

Subtitles are a textual representation of the audio track, usually just the dialog and usually in a language other than the audio track dialog, intended for foreign language audience.

C.2 Captions for the Hearing Impaired

Captions for the Hearing Impaired are a textual representation of the audio track, usually including all sounds, and usually in the same language as the audio track dialog, intended for hearing impaired audiences.

C.3 Text for the Visually Impaired

Text for the Visually Impaired is a textual description of visual elements of the content and usually in the same language as the audio track dialog, intended for visually impaired audiences.

C.4 Commentary

Commentary provides extra information about the associated content (e.g. Producer Commentary) usually in the same language as the audio track dialog.

C.5 Karaoke

Karaoke is a textual representation of songs' lyrics, usually in the same language as the associated song.

Annex D StereoImageTrackFileType (Normative)

The StereoImageTrackFileType shall represent stereoscopic content consisting of two Track Files, each consisting of monoscopic essence. The Track Files referenced by the Left Eye and Right Eye elements shall contain views associated with the left and right eyes of the viewer.

```
<xs:complexType name="StereoImageTrackFileType">
  <xs:complexContent>
    <xs:extension base="cpl:BaseResourceType">
      <xs:sequence>
        <xs:element name="LeftEye" type="cpl:TrackFileType" />
        <xs:element name="RightEye" type="cpl:TrackFileType" />
      </xs:sequence>
    </xs:extension>
  </xs:complexContent>
</xs:complexType>
```

The SourceDuration of an instance of StereoImageTrackFileType shall be equal to the corresponding parameters contained in its LeftEye and RightEye elements.

Annex E Sparse Data Essence (Normative)

In some situations, the data essence underlying a virtual track may be altogether absent from a given Segment, e.g. a Sequence within a Composition may not have any captions associated with it. In these situations, the Sequence associated with the virtual track within such a Segment should contain a single Resource and this Resource should be associated with an otherwise valid asset conveying the absence of essence for a duration equal to that of the Segment.

Figure E.1 depicts an example where no captions are associated with two Segments of a Composition. Each of these two Segments contains a single Sequence that itself contains a single Resource. These two Resources reference the same Track File ID=AF..12, which is a valid Data Essence Track File (an MXF File that wraps an XML representation of the caption timeline) that contains no actual caption (the timeline is empty).

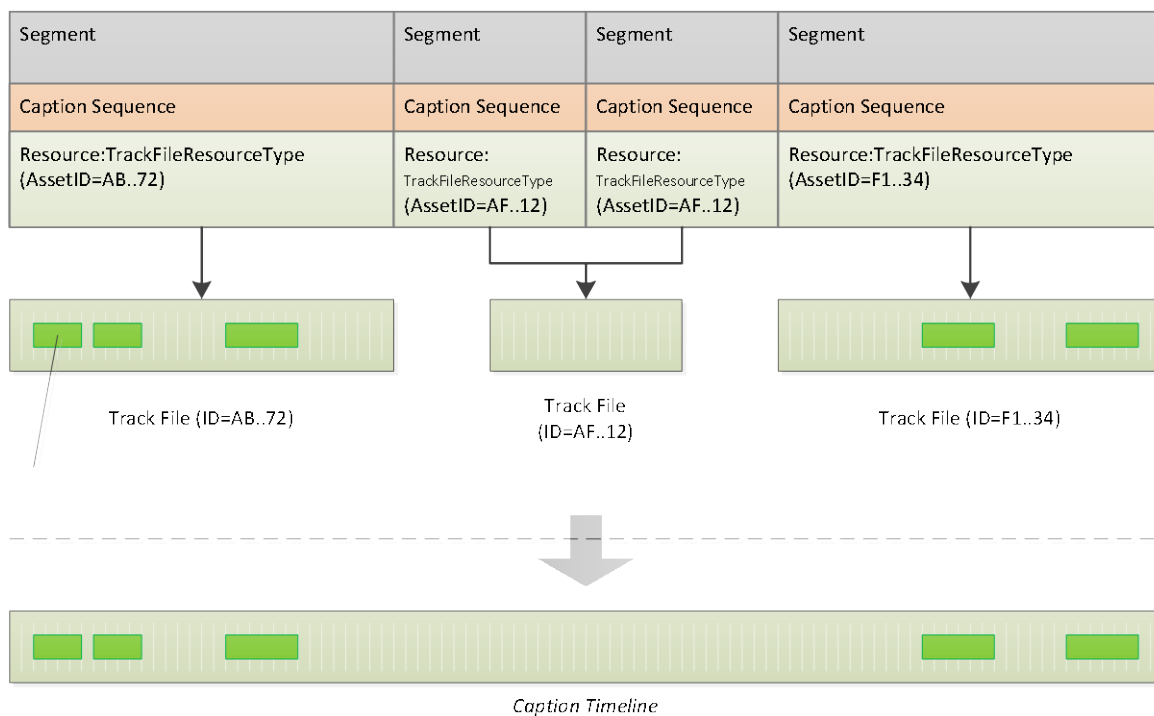


Figure E.1 – Sparse Data Essence

Annex F Additional Generic Sound Essence Descriptor Items (Normative)

F.1 General

Annex F shall specify items for the Generic Sound Essence Descriptor (see SMPTE ST 377-1).

F.2 Reference Image Edit Rate

Table F.1 – Specification of the Reference Image Edit Rate Item

<i>Item Name</i>	<i>Item Symbol</i>	<i>Type</i>	<i>Len</i>	<i>Local Tag</i>	<i>Item UL</i>	<i>Req ?</i>	<i>Meaning</i>	<i>Default</i>
Reference Image Edit Rate	ReferenceImageEditRate	Rational	8	dyn	06.0E.2B. 34.01.01. 01.0E.04. 02.01.01. 06.00.00. 00	Opt	Edit rate of the image essence with which the audio is intended to sync. For example, 24/1.001 fps, 24 fps, 25 fps	

The Reference Image Edit Rate shall indicate the edit rate of the image to which the audio is designed to sync. Absence of the item shall indicate that the Reference Image Edit Rate is unknown.

F.3 Reference Audio Alignment Level

Table F.2 – Specification of the Reference Audio Alignment Level Item

<i>Item Name</i>	<i>Item Symbol</i>	<i>Type</i>	<i>Len</i>	<i>Local Tag</i>	<i>Item UL</i>	<i>Req ?</i>	<i>Meaning</i>	<i>Default</i>
Reference Audio Alignment Level	ReferenceAudioAlignmentLevel	Int8	1	dyn	06.0E.2B. 34.01.01. 01.0E.04. 02.01.01. 07.00.00. 00	Opt	Reference alignment level of the audio expressed in dBFS. For example, -20dBFS.	

The Reference Audio Alignment Level shall be the reference alignment level of the audio expressed in dBFS. This is the number of dB below full scale on a PPM meter that represents the analog “0 VU” nominal alignment level, for example -20 dBFS. Absence of the item shall indicate that the Audio Reference Alignment Level is unknown.

Annex G Alternative Center Cuts (Normative)

The following specifies an additional optional item for the Generic Picture Essence Descriptor (see SMPTE ST 377-1).

Table shall specify the Alternative Center Cuts item.

The Alternative Center Cuts item specified in Table G.1 contains zero or more Alternative Center Cut values.

Table G.1 – Specification of the Alternative Center Cuts Item

<i>Item Name</i>	<i>Item Symbol</i>	<i>Type</i>	<i>Len</i>	<i>Local Tag</i>	<i>Item UL</i>	<i>Req ?</i>	<i>Meaning</i>	<i>Default</i>
Alternative Center Cuts	AlternativeCenterCuts	Batch of UL	Var	dyn	06.0E.2B. 34.01.01. 01.0E.04. 01.03.02. 0B.00.00. 00	Opt	Specifies the alternate aspect ratio subset(s) of the Active Area Rectangle, containing all of the critical action and centered on the Active Area Rectangle.	

As illustrated in Figure G.1, each Alternative Center Cut value shall be a UL defining an aspect ratio for a rectangular area containing all of the critical action that is (a) centered on the Active Area Rectangle, (b) entirely contained within the Active Area Rectangle and (c) has a height or width equal to that of the Active Area Rectangle. Table G.2 defines Alternative Center Cut values, which may be used as values in the Alternative Center Cuts batch. The Active Area Rectangle is defined in Annex H.

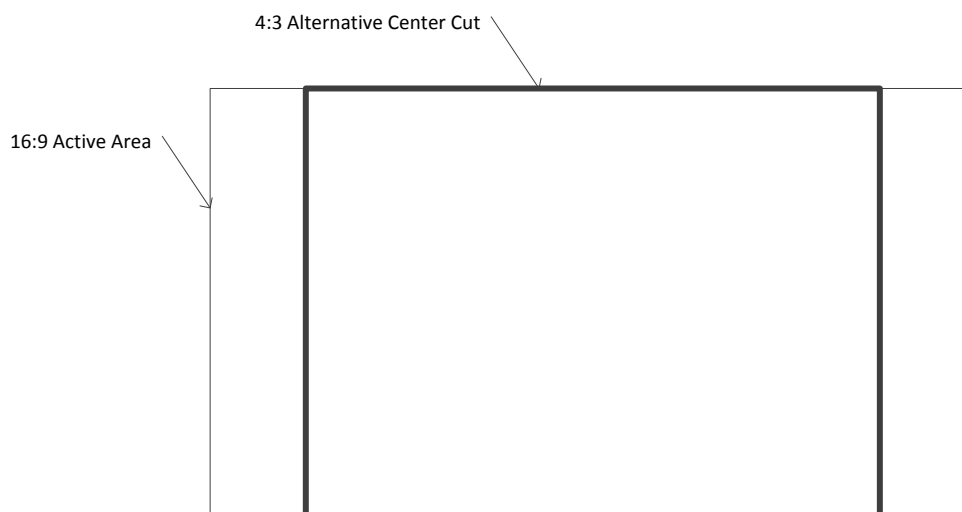


Figure G.1 – Alternative Center Cut Example

The presence of an Alternative Center Cut value shall not indicate that a cut is necessary, but merely that a cut is possible.

The presence of more than one Alternative Center Cut values shall indicate that more than one Alternative Center Cut are possible.

When generating an output deliverable from an IMP, an Alternative Center Cut value can be used to generate an AFD word (see SMPTE ST 2016-1) or crop the image to match the aspect ratio of the output. For example, given 16:9 image content associated with a 4:3 Alternative Center Cut value, a processing device can, depending on the downstream delivery requirements, output a 16:9 deliverable that contains an AFD word signaling a 4:3 Center Cut or create a 4:3 center cut deliverable by cropping the 16:9 image.

Table G.2 – Alternative Center Cut UL values

Byte No.	Name	Symbol	Value (hex)	Meaning
1-7	See Picture Source Characteristics as specified in SMPTE RP 224			
8	Version Number		0Dh	Registry Version at the point of registration of this label
9-12	See Picture Source Characteristics as specified in SMPTE RP 224			
13	Alternative Center Cut	AlternativeCenterCut	04h	Rectangular area of specified aspect ratio containing all of the critical action that is (a) centered on the Active Area Rectangle, (b) entirely contained within the Active Area Rectangle and (c) has a height or width equal to that of the Active Area Rectangle.
14	4:3 Alternative Center Cut	CenterCut43	01h	Indicates that the image essence can accommodate an alternative center cut with a 4:3 aspect ratio.
	14:9 Alternative Center Cut	CenterCut149	02h	Indicates that the image essence can accommodate an alternative center cut with a 14:9 aspect ratio.

Annex H Active Area Rectangle (Normative)

This section specifies additional optional items for the Generic Picture Essence Descriptor (see SMPTE ST 377-1)

The items specified in Table H.1 shall define the Active Area Rectangle.

The Active Area Rectangle shall be the rectangular region which is intended to be visible to the viewer, at the sole discretion of the author. As such the Active Area Rectangle may contain letterboxing or side mattes.

As illustrated in Figure H.1, the coordinates of the Active Area Rectangle are defined relative to the Display Rectangle (see SMPTE ST 377-1).

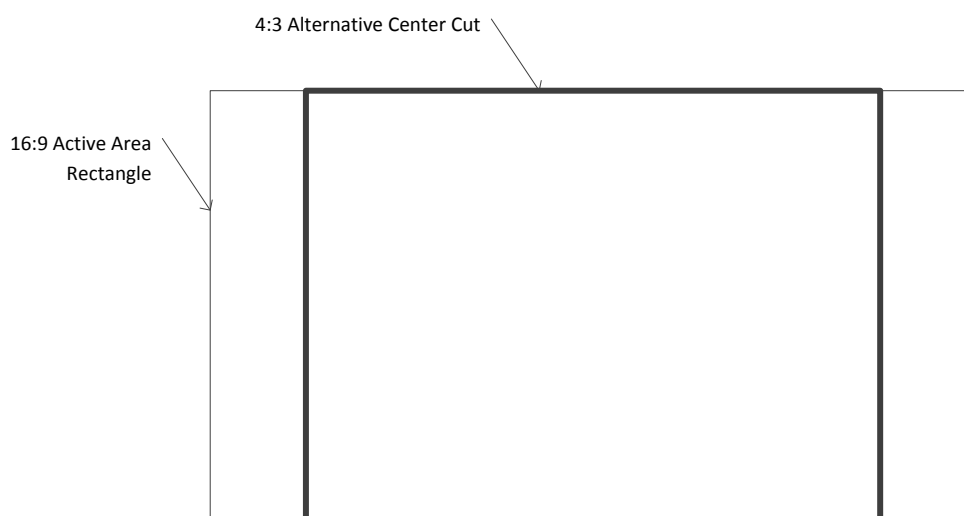


Figure H.1 – Active Area Rectangle

The Active Area Rectangle shall be defined by its Width and Height Properties in pixels. These Properties shall default to the same value as the Display Width and Display Height. The Active Rectangle Width and Height values shall be no greater than the Display Width and Display Height values taking into account any Active X Offset and Active Y Offset Property values.

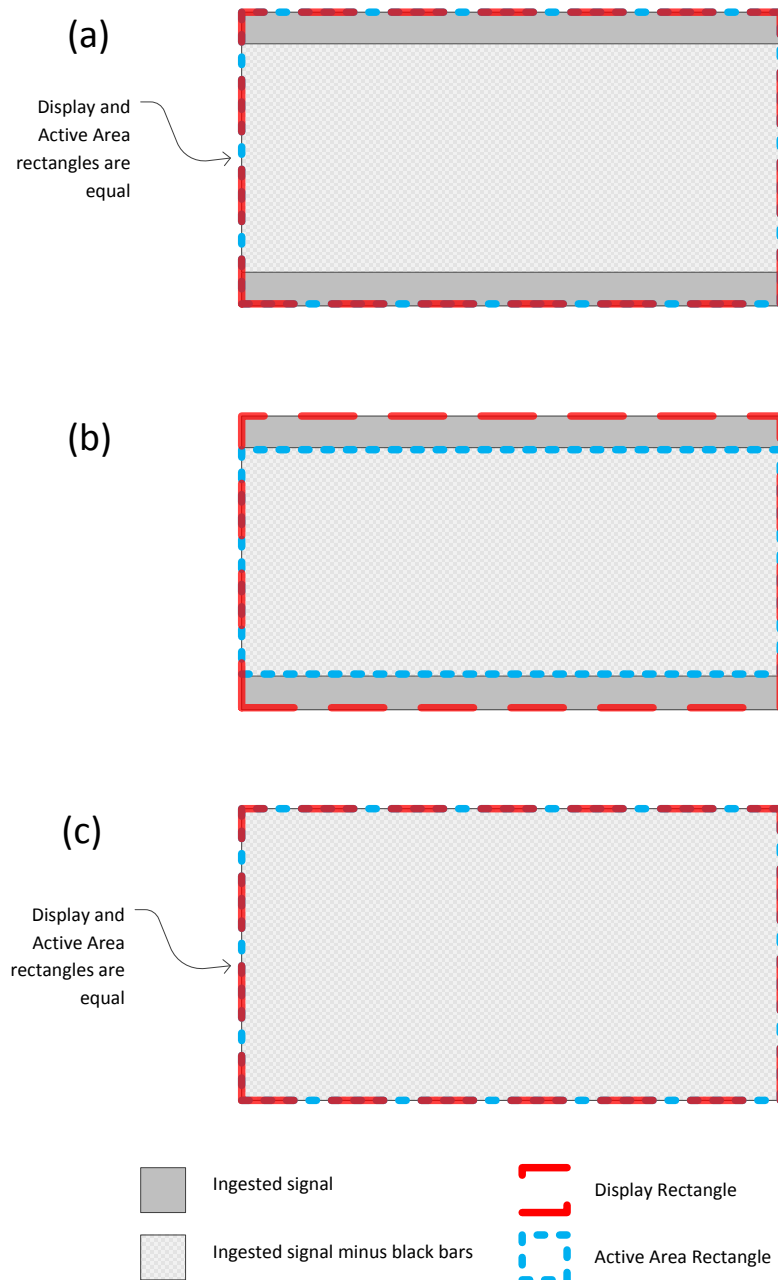
The mapping from the Display Rectangle to the Active Area Rectangle shall be defined by the Active X Offset and Active Y Offset Properties, which give the zero-based coordinates of the Active Area Rectangle relative to the upper left corner of the Display Rectangle. Their values shall not be negative.

Table H.1 – Specification of the Active Area Rectangle

<i>Item Name</i>	<i>Item Symbol</i>	<i>Type</i>	<i>Len</i>	<i>Local Tag</i>	<i>Item UL</i>	<i>Req ?</i>	<i>Meaning</i>	<i>Default</i>
Active Height	ActiveHeight	UInt32	4	dyn	06.0E.2B.34 .01.01.01.0 E.04.01.05. 01.13.00.00. 00	Opt	Specifies the height of the Active Area Rectangle.	Display Height
Active Width	ActiveWidth	UInt32	4	dyn	06.0E.2B.34 .01.01.01.0 E.04.01.05. 01.14.00.00. 00	Opt	Specifies the width of the Active Area Rectangle.	Display Width
Active X Offset	ActiveXOffset	UInt32	4	dyn	06.0E.2B.34 .01.01.01.0 E.04.01.05. 01.15.00.00. 00	Opt	Specifies the horizontal offset of the Active Area Rectangle from the left side of the Display Rectangle.	0
Active Y Offset	ActiveYOffset	UInt32	4	dyn	06.0E.2B.34 .01.01.01.0 E.04.01.05. 01.16.00.00. 00	Opt	Specifies the vertical offset of the Active Area Rectangle from the upper side of the Display Rectangle.	0

	<i>a</i>	<i>b</i>	<i>c</i>
Display Rectangle	1920x1080	1920x1080	1920x1080
Active Area Rectangle	1920x1080	1920x800	1920x1080
DisplayHeight	1080	1080	1080
DisplayWidth	1920	1920	1920
ActiveHeight	1080	800	1080
ActiveWidth	1920	1920	1920
ActiveXOffset	0	0	0
ActiveYOffset	0	140	0

Figure H.2 illustrates examples of the use of the Active Area Rectangle. For instance, as shown in (a) and (b), two different users may elect to identify different portions of the same image as the Active Area Rectangle.



	<i>a</i>	<i>b</i>	<i>c</i>
Display Rectangle	1920x1080	1920x1080	1920x1080
Active Area Rectangle	1920x1080	1920x800	1920x1080
DisplayHeight	1080	1080	1080
DisplayWidth	1920	1920	1920
ActiveHeight	1080	800	1080
ActiveWidth	1920	1920	1920
ActiveXOffset	0	0	0
ActiveYOffset	0	140	0

Figure H.2 – Display Rectangle and Active Area Rectangle Scenarios (Informative)

Annex I RFC 5646 Language Tag List (Normative)

The following specifies an additional optional item for the Timed Text Descriptor (see SMPTE ST 429-5).

Table I.1 shall define the RFC 5646 Language Tag List item.

The RFC 5646 Language Tag List item specified in Table I.1 shall contain a list of one or more Language Tags (as specified in IETF RFC 5646) using the following syntax.

`RFC5646LanguageTagList = Language-Tag *("," Language-Tag)`

Table I.1 – Specification of the RFC 5646 Language Tag List Item

<i>Item Name</i>	<i>Item Symbol</i>	<i>Type</i>	<i>Len</i>	<i>Local Tag</i>	<i>Item UL</i>	<i>Req ?</i>	<i>Meaning</i>
RFC 5646 Language Tag List	RFC5646LanguageTagList	UTF16	Var	dyn	06.0E.2B. 34.01.01. 01.0E.03. 01.01.02. 02.16.00. 00	Opt	Identifies one or more languages as a comma-separated list of RFC 5646 language tags.