

SMPTE STANDARD

Interoperable Master Format — Application #2E



Page 1 of 34

Table of Contents	Page
Foreword	2
Intellectual Property	2
1 Scope	3
2 Conformance Notation	3
3 Normative References	3
4 Overall	4
5 Image Essence	5
6 Image Track Files	11
7 Composition	17
Annex A ITU-R BT.2020 Transfer Characteristic Label (Normative)	19
Annex B Mastering Display Color Volume Metadata (Normative)	20
Annex C SMPTE ST 2084 Transfer Characteristic Label (Normative)	23
Annex D ITU-R BT.2020 Non-Constant Luminance Coding Equations Label (Normative)	24
Annex E Consolidated Schema (Informative)	25
Annex F Recommended selection of JPEG 2000 Profiles (Informative)	26
Annex G Additional JPEG 2000 Picture Essence Compression Labels (Normative)	29
Annex H Image Frame And Active Area Rectangle Examples (Informative)	30
Bibliography (Informative)	34

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 its Standards Operations Manual. This SMPTE Engineering Document 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.

1 Scope

This specification defines IMF Application #2E. It is a specialization of the IMF Framework.

Application #2E is meant for studio applications where a TV or movie title is transformed into multiple content versions (airline edits, special edition, languages...) that are made available to multiple consumer distribution channels (Internet, optical media, broadcast...) across multiple territories and over the span of many months to over a year. It uses image essence coded as a JPEG 2000 codestream and audio essence coded as linear PCM.

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; then formal languages; then figures; and then any other language forms.

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

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

SMPTE ST 422:2014, Material Exchange Format — Mapping JPEG 2000 Codestreams into the MXF Generic Container

SMPTE ST 2067-2:2020, Interoperable Master Format – Core Constraints

SMPTE ST 2084:2014, High Dynamic Range Electro-Optical Transfer Function of Mastering Reference Displays

SMPTE ST 2086:2018, Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images

SMPTE ST 2113:2018, Colorimetry of P3 Color Spaces

Recommendation ITU-R BT.601-7, Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios

Recommendation ITU-R BT.709-6, Parameter Values for the HDTV Standards for Production and International Programme Exchange

Recommendation ITU-R BT.2020-2, Parameter Values for Ultra-High Definition Television Systems for Production and International Programme Exchange

CTA 861-G, A DTV Profile for Uncompressed High Speed Digital Interfaces (November 2016)

IEC 61966-2-4 Edition 1.0, Multimedia Systems and Equipment – Colour Measurement and Management — Part 2-4: Colour Management — Extended-Gamut YCC Colour Space for Video Applications — xvYCC

ISO/IEC 15444-1:2016, Information Technology — JPEG 2000 Image Coding System: Core Coding System

ISO 11664-3:2012 (CIE S014-3/E:2011), Colorimetry — Part 3: CIE Tristimulus Values

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

4 Overall

4.1 General

All provisions of SMPTE ST 2067-2 shall apply.

4.2 Format

Track Files shall conform to SMPTE ST 379-1.

4.3 Shim Parameters

Table 1 – Shim Parameter Values Definitions

<i>Shim Parameter</i>	<i>Value</i>
shim_id	http://www.smpte-ra.org/schemas/2067-21/2016
picture_bitrate	ST 2067-21
picture_format	ST 2067-21
gc_type	379-1-gc
picture_family	JPEG2000
picture_custom_ANC	false
picture_render_ANC	false

4.4 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 2.

Table 2 – XML Schema root element definition

```
<xs:schema targetNamespace="http://www.smpte-ra.org/schemas/2067-21/2016"
  xmlns:app2e="http://www.smpte-ra.org/schemas/2067-21/2016"
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  elementFormDefault="qualified" attributeFormDefault="unqualified">

  <!-- schema definitions found in this document -->

</xs:schema>
```

5 Image Essence

5.1 General

Image essence shall consist of image frames, each a rectangular pixel array.

5.2 Constraints

Each image frame shall conform to one of the permitted combinations of image frame characteristics listed in Table 3.

The notation *a..b* indicates that any value between *a* and *b*, including *a* and *b*, is allowed. For instance, the range 1..4096 includes the value 3840 and the range 1..3112 includes the value 2160.

Table 3 – Image Characteristics

Characteristic	Constraint								
Image Frame Width	1..1920		1..3840			1..4096			
Image Frame Height	1..1080		1..2160			1..3112			
Colorimetry	COLOR.1 COLOR.2 COLOR.3		COLOR.4	COLOR.3	COLOR.5	COLOR.7	COLOR.3	COLOR.5	COLOR.6 COLOR.7
Pixel Bit Depth	8 10			8 10 12 16	10 12	10 12 16	8 10 12 16	10 12	10 12 16
Frame Structure	Inter-laced	Progressive							
Stereoscopy	Mono-scopic	Stereoscopic Monoscopic							
Frame Rate	25 30 30000/ 1001	24 24000/1001 25 30 30000/1001 50 60 60000/1001				24 24000/1001 25 30 30000/1001 50 60 60000/1001 120			
Sampling	4:2:2		4:4:4	4:2:2			4:4:4		
Quantization	QE.1						QE.1 QE.2		
Color Components	Y'C'B'C'R	Y'C'B'C'R R'G'B'	Y'C'B'C'R				R'G'B'		

EXAMPLE 1: An image frame that combines COLOR.6 colorimetry with Y'C'B'C'R color components is not supported.

EXAMPLE 2: A monoscopic progressive R'G'B' 4:4:4 image frame with dimensions 3840x2160 that combines COLOR.7 colorimetry with 12-bit pixel bit depth, 60 Hz frame rate, and QE.2 quantization is supported.

Note 1: This specification does not support the Y'c'BcC'RC (constant luminance) color components specified in Recommendation ITU-R BT.2020.

Note 2: IEC 61966-2-4 uses Y'C'B'C'R and YCC (luma-chroma-chroma) interchangeably.

Note 3: Image formats defined in this document do not necessarily correspond to image formats defined in other standards or recommendations.

5.2.1 Frame Dimensions

The width and height of the frame are defined as the number of horizontal and vertical pixels, respectively.

5.2.2 Frame Structure

5.2.2.1 General

Implementations shall support image frames with either a progressive or interlaced structure.

5.2.2.2 Progressive Structure

An image frame with progressive structure shall consist of a complete image frame, scanned progressively left to right and from top to bottom.

5.2.2.3 Interlaced Structure

An image frame with interlaced structure shall consist of a pair of fields, a first field then a second field. The lines of each field shall have twice the vertical spatial sampling pitch of the frame. Lines in the second field shall be displaced vertically by the vertical sampling pitch and the line timing shall be delayed temporally by half the frame time from the lines in the first field.

The temporal order and relative line positions of the two fields are indicated by the Field Dominance and DisplayF2Offset items described in Section 6.2.1, Table 8.

5.2.2.4 Frame Size

The height and width of the image frame shall be an integer.

Image frames with an interlaced image structure shall have an even number of vertical pixels.

5.2.2.5 Frame Rate

When interlaced frame structure is used, the field rate, i.e. the number of image fields per second, shall be twice the frame rate.

5.2.2.6 Color Components

Implementations shall support image frames sampled using either R'G'B' or Y'C_BC_R color component triplets.

5.2.2.7 Pixel Bit Depth

Implementations shall support the color component of each pixel being represented by an integer in the set $\{0..2^n - 1\}$, with n being the pixel bit depth.

5.2.2.8 Sampling

In 4:4:4 sampling, each component shall be sampled once at each image frame pixel.

In 4:2:2 Y'C_BC_R sampling, the Y' component shall be sampled at each pixel, but the C_R and C_B components shall be horizontally subsampled by a factor of two with respect to the Y component, co-sited with even-numbered Y' samples.

5.2.2.9 Stereoscopic and Monoscopic Image Essence

Monoscopic essence consists of a single sequence of image frames.

Stereoscopic essence consists of a sequence of pairs of image frames, a left eye frame and a right eye frame, for stereoscopic viewing. The two images of a pair shall be coincident in time.

5.3 Colorimetry

Implementations shall support the mappings of component signals to red, green and blue tristimulus values listed in Table 4.

Table 4 – Colorimetry Systems

System	Description
COLOR.1	Mapped as specified for 625-line systems in Section 2.6 of Recommendation ITU-R BT.601-7.
COLOR.2	Mapped as specified for 525-line systems in Section 2.6 of Recommendation ITU-R BT.601-7.
COLOR.3	Mapped as specified in Section 1 of Recommendation ITU-R BT.709-6.
COLOR.4	Mapped using method xvYCC709 as specified in IEC 61966-2-4.
COLOR.5	Mapped as specified in Recommendation ITU-R BT.2020-2. R'G'B' components are mapped to Y'C _B C _R components using the (non-constant luminance) derivation of Y' and color difference signals specified in Table 4 of Recommendation ITU-R BT.2020-2.
COLOR.6	Mapped using the P3D65 color system color primaries and white point specified in SMPTE ST 2113 and the transfer function specified in SMPTE ST 2084.
COLOR.7	R'G'B' components are mapped using the color primaries and white point specified in Recommendation ITU-R BT.2020 and the transfer function specified in SMPTE ST 2084. R'G'B' components are mapped to Y'C _B C _R components using the (non-constant luminance) derivation of Y' and color difference signals specified in Table 4 of Recommendation ITU-R BT.2020-2.

Note: In Recommendation ITU-R BT.601 and Recommendation ITU-R BT.709, the signals R', G', B', Y', C_B' and C_R' are referred to as signals E_R', E_G', E_B', E_Y', E_{CB}' and E_{CR}' respectively, i.e. they correspond to gamma pre-corrected signals.

5.4 Quantization

Implementations shall support R'G'B' or Y'C_BC_R component signals being quantized according to one of the systems specified in Table 5.

Table 5 – Quantization Systems

System	Component Triplet	Quantization equations (<i>n</i> is the pixel bit depth)	Notes
QE.1	R'G'B' Y'C _B C _R	$D'_R = \text{INT}((219 \cdot R' + 16) \cdot 2^{n-8})$	Equivalent to quantization equations of Section 3 of Recommendation ITU-R BT.709-6, and Table 3 and Table 4 of Recommendation ITU-R BT.601-7.
		$D'_G = \text{INT}((219 \cdot G' + 16) \cdot 2^{n-8})$	
		$D'_B = \text{INT}((219 \cdot B' + 16) \cdot 2^{n-8})$	
		$D'_Y = \text{INT}((219 \cdot Y' + 16) \cdot 2^{n-8})$	
		$D'_{CR} = \text{INT}((224 \cdot C'_R + 128) \cdot 2^{n-8})$	
		$D'_{CB} = \text{INT}((224 \cdot C'_B + 128) \cdot 2^{n-8})$	
QE.2	R'G'B'	$D'_R = \text{INT}(R' \cdot (2^n - 1))$	The mapping of components signals using the QE.2 system onto interfaces such as HD-SDI is defined in other specifications.
		$D'_G = \text{INT}(G' \cdot (2^n - 1))$	
		$D'_B = \text{INT}(B' \cdot (2^n - 1))$	

5.5 Encoding Profile

5.5.1 Single Codestream

Each frame, in the case of progressive structure, or field, in the case of interlaced structure, shall be encoded as a single codestream.

5.5.2 JPEG 2000 Profiles

Implementations shall support the combinations of JPEG 2000 profiles (as specified in ISO/IEC 15444-1:2016) and image frame dimensions listed in Table 6.

Note: See Section 5.2 for a definition of the notation *a...b*.

Table 6 – JPEG 2000 Profiles

Image Frame Width	1..3840		1..2048		2049..4096	
Image Frame Height	1..2160		1..1556		1..3112	
JPEG 2000 Profile	Broadcast Contribution Single Tile Profile	Broadcast Contribution Multi-tile Reversible Profile	2k IMF single/multi-tile reversible profile	2k IMF Single Tile Lossy Profile	4k IMF single/multi-tile reversible profile	4k IMF Single Tile Lossy Profile
JPEG 2000 Operating Levels	Level 1 Level 2 Level 3 Level 4 Level 5	Level 6 Level 7	Mainlevel 1 Mainlevel 2 Mainlevel 3 Mainlevel 4 Mainlevel 5 Mainlevel 6		Mainlevel 1 Mainlevel 2 Mainlevel 3 Mainlevel 4 Mainlevel 5 Mainlevel 6 Mainlevel 7 Mainlevel 8	
JPEG 2000 Operating Sublevels	n/a		Sublevel 0 only	All allowed at a given Mainlevel with the exception of Sublevel 0	Sublevel 0 only	All allowed at a given Mainlevel with the exception of Sublevel 0

Note 1: The JPEG 2000 Broadcast Contribution profiles do not support either 16-bit Pixel Bit Depth or a Maximum Components Sampling Rate greater than $520 \cdot 10^6$ samples s^{-1} . In both cases, JPEG 2000 IMF profiles are used instead.

Note 2: The JPEG 2000 profiles, operating level and sublevel in Table 6 are selected as described in Annex F.

5.5.3 Component Ordering

In a codestream, color components shall be ordered as specified in Table 7.

Table 7 – JPEG 2000 Color Component Ordering

Component Index	$R'G'B'$	$Y'C'_BC'_R$
0	R'	Y'
1	G'	C'B
2	B'	C'R

6 Image Track Files

6.1 Essence

Image Track Files shall contain image essence conforming to Section 5.

6.1.1 Wrapping

Image Track Files shall conform to SMPTE ST 422.

In the case of progressive frame structure, the image essence shall be wrapped according to mode P1 specified in SMPTE ST 422 ("Frame-wrapping").

In the case of interlaced frame structure, the image essence shall be wrapped according to mode I1 specified in SMPTE ST 422 ("Interlaced Frame, 1 field per KLV Element").

The Top-Level File Package of Image Track File shall reference:

- a CDCI Picture Essence Descriptor SMPTE ST 377-1 if the image uses $Y'C'_BC'_R$ color components; or
- an RGBA Picture Essence Descriptor SMPTE ST 377-1 if the image essence uses $R'G'B'$ color components.

6.2 Generic Picture Essence Descriptor

6.2.1 General

The Generic Picture Essence Descriptor items (including those specified in SMPTE ST 2067-2) shall be constrained as specified in Table 8.

Table 8 – Generic Picture Essence Descriptor Items

Generic Picture Essence Descriptor Item	Constraints
Sample Rate	See Annex A of SMPTE ST 422:2014.
Signal Standard	Shall be ignored.
Frame Layout	See Section 6.2.1.3.
Stored Width	See Section 6.2.1.1.
Stored Height	See Section 6.2.1.1.
StoredF2Offset	Shall not be present.
Sampled Width	Shall not be present or shall be equal to Stored Width.
Sampled Height	Shall not be present or shall be equal to Stored Height.
SampledXOffset	Shall not be present or shall be 0.
SampledYOffset	Shall not be present or shall be 0.